AD-A077 994 DESMATICS INC STATE COLLEGE PA F/G 12/1
AN APPROACH FOR PROVIDING MORE ACCURATE PROBABILITY ASSESSMENTS--ETC(U)
MAY 78 D E SMITH , R L GARDNER DAHC19-76-C-0045
UNCLASSIFIED ARI-TR-78-B7 NL

1 OF 2
AD-
A077994

LEVEL IV

ARI TECHNICAL REPORT
TR-78-B7

AD A077994

AN APPROACH FOR PROVIDING MORE ACCURATE PROBABILITY ASSESSMENTS

by

Dennis E. Smith and Robert L. Gardner
Desmatics, Inc.
P.O. Box 618
State College, Pa. 16801

May 1978

Contract DAHC 19-76-C-0045 new

Monitored Technically by S. M. Halpin, Battlefield
Information Systems Technical Area, ARI

Prepared for

U.S. ARMY RESEARCH INSTITUTE
for the BEHAVIORAL and SOCIAL SCIENCES
5001 Eisenhower Avenue
Alexandria, Virginia 22333

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

JOSEPH ZEIDNER Acting Technical Director	W. C. MAUS COL, GS Commander

Research accomplished
under contract to the Department of the Army

Desmatics, Inc.

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U. S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-P, 5001 Eisenhower Avenue, Alexandria, Virginia 22333.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U. S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

(18) (9)

ARI TECHNICAL REPORT

(19) TR-78-B7

(6)

AN APPROACH FOR PROVIDING MORE ACCURATE PROBABILITY ASSESSMENTS.

by

(10) Dennis E. Smith and Robert L. Gardner
Desmatics, Inc.
P.O. Box 618
State College, Pa. 16801

(11) May 1978 (12) 104

(15) Contract DAHC 19-76-C-0045

Monitored Technically by S. M. Halpin, Battlefield Information Systems Technical Area, ARI

Accession For	
NTIS GRA&I	☒
DDC TAB	☐
Unannounced	☐
Justification ______	
By ______	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

Prepared for

U.S. ARMY RESEARCH INSTITUTE
for the BEHAVIORAL and SOCIAL SCIENCES
5001 Eisenhower Avenue
Alexandria, Virginia 22333

Army Project Number (16) 2Q761102B74F — Intelligence Information

This is a report on an exploratory research project designed to meet military management requirements for research on a specific management problem. A limited distribution is made, primarily to the operating agencies directly involved.

391 156 LB

ABSTRACT

This report describes a possible approach for providing more accurate probability assessments in those real-world decision-making situations for which relative frequency information generally does not exist. In this approach an estimated functional relationship is obtained from a comparison, in relative frequency base (RFB) problems, of a person's posterior probability assessments with the true posterior probabilities. Then, the functional relationship is used as a model to provide probability predictions for new problems by adjusting the person's corresponding probability assessments. Experimental evidence indicates that the resulting models can provide satisfactory probability predictions for problems having identical structure to those used in model construction. At this time, some obstacles prevent immediate use of these models for probability prediction in non-relative frequency base (NRFB) problems. However, it would be unwise to discount the feasibility of developing a practical RFB-based model for adjusting real-world probability estimates until the approach were tested with participants and problems in a field of common expertise (such as intelligence analysis or weather forecasting, for example) where more uniformity of background and experience exists.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

I. INTRODUCTION

Decision-making is usually a complex and difficult task. In many cases, however, this task may be greatly simplified by the application of numerical techniques, particularly if probabilities are used to express relationships and uncertainties within the structure of the underlying problem. Thus, probability estimation (or probability assessment[1]) is a valuable step in the decision-making process.

The intelligence analyst, for example, is often called upon to implicitly or explicitly estimate probabilities. Traditionally, he has reflected his judgments about "raw" intelligence according to source reliability and content vali ty using a letter scale of A through F for source reliability and a numerical scale of 1 through 6 for content validity. In addition, or as a substitute, the analyst has expressed his judgments in verbal descriptions such as "it is doubtful that", "chances are good that", "it is fairly certain that", etc. Unfortunately, verbal descriptions of the uncertain link between data and predictions may tend to confuse rather than clarify. This contention was borne out in a study by Johnson (1973), who found experimental indications that "...the use of qualitative expressions to communicate the accuracy or relative likelihood of occurrences of intelligence data and products will often result in a high degree of misunderstanding."

---

[1] Brown, Kahr, and Peterson (1974) use the term "probability assessment" when the true probability is unknown, reserving the term "probability estimation" for those situations in which the true probability may be determined. This distinction is not being made in this report; "estimation" and "assessment" are used interchangeably.

Numerical probability statements (e.g., "I believe there is a 70% chance that event E will occur.") may well provide a more understandable means of communicating. As noted by Brown and Shuford (1973), the use of numerical probabilities to express uncertainty provides concise, relevant quantitative data. In general, however, it is extremely difficult to assess the calibration (i.e., the accuracy) of numerical probability estimates unless relative frequency information exists. Unfortunately, such information is not available in most real-world decision-making situations.

The fundamental tenet of the research described in this report is that the performance of persons making probability estimates in problems having a <u>relative frequency base</u> (RFB) can yield information with which to calibrate their estimates made in situations lacking a relative frequency base. *The reason for using RFB situations is that*, unlike <u>non-relative frequency base</u> (NRFB) problems, probability estimates made in RFB problems can be compared with a normative Bayesian model to derive models of each individual's inherent estimation biases and errors. These functional models can then be used to adjust the estimates made in NRFB probability estimation tasks.

II. BACKGROUND

In the decision-making framework, expression of uncertainty by means of numerical probabilities involves not only probability estimation, but also the attendant task of updating probabilities in light of new information. Most experiments to investigate probability estimation and updating have involved "book-bag and poker chip" problems. The following description of these problems is given by Winkler and Murphy (1973).

> A commonly used device in probability revision experiments is the bookbag-and-poker-chip paradigm, in which the experimenter has a number (usually two) of bookbags filled with poker chips. The composition of the bookbags differs with respect to the number of poker chips of a particular color; the first bookbag may contain 70 red chips and 30 blue chips, for example, whereas the second bookbag contains 30 red chips and 70 blue chips. The subject is told that one of the bookbags has been chosen by a random device, and his task is to assess probabilities for the possible bookbags. Since the subject is told that the bookbag has been selected randomly, the initial, or prior, probabilities should be equal. The subject is then given information in the form of a series of draws, at random and with replacement, from the chosen bookbag. Upon observing the results of these draws, the subject assesses his posterior probabilities for the possible bookbags.

Thus, these problems have a relative frequency *base* (*RFB*) *as do the* "balls and urns" problems often used in introductory probability courses to illustrate the application of Bayes' theorem.

A. BAYES' THEOREM AND DISCORDANCE

Bayes' theorem serves as an optimal normative model for updating

probability estimates. The usual form of this theorem is

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{\sum_k P(D|H_k)P(H_k)}$$

where $P(H_i)$ denotes the prior probability of hypothesis H_i, $P(D|H_i)$ denotes the conditional probability of specific data D occurring given that hypothesis H_i is true, and $P(H_i|D)$ denotes the posterior probability of hypothesis H_i given the occurrence of the data D.

Numerous studies[1] have indicated what has been called "conservatism" in estimation of $P(H_i|D)$, posterior probability. This terminology was used because most persons tend to revise their estimates from the prior probability by an amount which is <u>less</u> than that indicated by Bayes' theorem. The more general term "discordance" will be used in this report to reflect <u>any</u> difference between an estimated probability and the true probability.

Much research has been aimed at determining how departures from the normative Bayesian model are affected by various aspects of the experimental situation, and at developing theories to account for discordant behavior. In general, the emphasis has been on conservatism, which has been found to depend on a number of factors, such as the response mode, the diagnosticity of the data, the perceived prior probabilities, the length of the data sequence presented, and the composition of the data.

In RFB problems estimates of $P(D|H_i)$ are usually accurately and

---

[1] See, for example, Edwards (1966b), Phillips, Hays, and Edwards (1966), Wheeler and Beach (1968), Chinnis and Peterson (1970), and Goodman (1972).

easily made. Therefore, the true value of $P(H_i|D)$ can be obtained without difficulty by application of Bayes' theorem. However, in the real world direct estimates of $P(H_i|D)$ cannot be based on a direct relative frequency assessment. Hence, estimation of posterior probabilities is a more complex task. It should be stressed that in most real-world (e.g., intelligence) situations neither $P(H_i|D)$ nor $P(D|H_i)$ has a relative frequency basis. Thus, both these types of probabilities are difficult to estimate, and it is quite likely that both are inaccurately estimated. If this, in fact, is the true situation, estimated posterior probabilities would usually be incorrect regardless of whether they were obtained directly by estimates or indirectly by Bayes' theorem.

Therefore, discordance is to be expected in estimates of $P(D|H_i)$. Then, despite the use of any computer aids, the value of $P(H_i|D)$ arrived at by using inaccurate estimates of $P(D|H_i)$ in conjunction with Bayes' theorem may be no more accurate than a direct estimate of that posterior probability. Thus, methods of evaluating and correcting for discordance in probability estimates are required in order to improve the quality of those estimates.

B. RESEARCH OVERVIEW

The research described in this report is aimed at development of an algorithm for predicting and correcting discordance in those real-world decision-making situations for which relative frequency information linking data with hypotheses generally will not exist. This research has concentrated on the simplest type of problem, the two-hypothesis problem with equal prior probabilities.

The rationale underlying the decision to deal with this type of problem rests on the assumption that if the suggested algorithm does not prove promising when based on uncomplicated examples, there would be little hope that it would prove promising on more complex (and realistic) problems of the type found in Army tactical decision-making Furthermore, by restricting attention to simple problems, the participating subjects would be able to complete more problems within a given period of time. Hence, more data would be available for estimating the required functions and for testing the validity of the suggested algorithm. Thus, at this stage of the research, the use of a number of short straightforward problems would appear to offer a higher payoff than the use of a smaller number of more complex (and perhaps more realistic) problems.

Unlike the "book-bag and poker chip" problems, those problems lacking a relative frequency base offer no direct method for assessing the differences between estimated and true probabilities, either for $P(H_i|D)$ or for $P(D|H_i)$. Because of this, the general procedure must be an indirect one.

In designing an experiment for observing behavior in RFB problems, consideration must be given to its fundamental purpose. If the objective were to obtain estimates which match as closely as possible those from the normative Bayesian model, then the experiment should employ the techniques which prior research has shown to reduce discordance. However, the goal here is to observe probability estimation in RFB situations in order to develop models which can be used in calibrating probability estimates made in NRFB situations. As a consequence, the present experiment allowed the responses elicited for RFB problems to exhibit the individual biases and errors inherent in each person's approach to probability estimation.

For a more rigorous statement of the hypothesis underlying the research, suppose that I people are involved in a probability estimation task. It is reasonable to assume an underlying functional relationship of the generic form

$$p = f_i(r) + \varepsilon$$

where:

- p denotes the true probability value
- r denotes an estimated probability value
- f_i denotes the functional relationship between p and r for person i
- ε denotes an error term resulting from random variation present in a person's probability estimates.

It should be noted that although probability values are assumed here for illustrative purposes, one or both of the variables p and r could be some <u>transform</u> of probability. The functional relationship f_i is somewhat akin to the "realism line" discussed by Sibley (1974). The realism line, which may be constructed from a person's probability estimates of confirmable future events, is a plot of observed relative frequency vs. estimated probability.

The realism line concept is related to the idea of a scoring rule. As described by Brown, Kahr and Peterson (1974), a scoring rule is a device for providing feedback to an estimator so that he may be rewarded in terms of how well his forecasts turn out. A <u>proper</u> scoring rule is one by which the assessor can maximize his score only by reporting his actual beliefs, since any other estimation or reporting strategy will

produce a lower score. A realism line derived for a subject can be evaluated in terms of a proper scoring rule. Conversely, application of a proper scoring rule can be used to generate a realism line from a set of observations.

Unfortunately the use of scoring rules and realism lines as decision aids has several practical disadvantages, the foremost of which is that one must obtain assessments for a large number of confirmable events and await their confirmation before the assessor can be informed of the results. This further implies that the assessor must be available first to make the assessments and also later, after the events have been confirmed, to receive the feedback. Another disadvantage is that any form of feedback or use of scoring rules may interfere with the assessment process and thus may inhibit performance rather than enhance it. The approach adopted in the current research obviates these difficulties by avoiding the use of confirmable events where the ultimate probabilities are either zero or one, substituting instead the use of problem situations involving probabilities covering the full (0,1) range.

In the current study, an estimated functional relationship $\hat{f}_i$ (not necessarily of the same form for each person) was obtained, via regression techniques, from a comparison of each person's posterior probability assessments in RFB problems with the true posterior probabilities. Then, using the posterior probability estimate r made by person i, $\hat{f}_i$ should provide a prediction (i.e., a corrected estimate) $\hat{p}$ of the true posterior probability. In mathematical notation, $\hat{p} = \hat{f}_i(r)$.

By looking at these functions for several people, it is possible to test whether or not the discordance observed in the estimates of $P(H|D)$

in a non-relative frequency base (NRFB) problem may be corrected. Given K NRFB problems in which posterior probabilities are to be estimated, a probability assessment r_{ik} may be obtained for each person and problem. If the degree of discordance is the same as in the RFB problems, the predicted "true" probabilities $\hat{p}_{ik}$ for $i = 1,\ldots,I$ [where $\hat{p}_{ik} = \hat{f}_i(r_{ik})$] should be unbiased estimates of the same quantity. That is, the estimated functional relationships $\hat{f}_i$ should provide a means by which assessments of NRFB probabilities may be calibrated.

The statistical procedures for testing the feasibility of correcting for discordance are described in Appendix A. That appendix includes a discussion of functional relationships, hypothesis testing, and statistical power considerations.

III. THE EXPERIMENT

If the degree of discordance in the NRFB problems were found to be the same as in the RFB problems, significant implications would exist for potential application in Army tactical decision-making. For example, the functions $\hat{f}_i$ could be used directly as a feedback mechanism with which to train probability estimators to improve their accuracy. As an alternative, the functions could instead be used as a correction device to minimize bias in probability estimates. This latter application, it should be noted, would not interfere with an analyst's method of estimating probabilities nor would it require that probability estimators receive special training to help them change their probability assessments.

To test the feasibility of correcting for discordance, an experiment was performed following the design discussed in Appendix A. Participants in the experiment were Pennsylvania State University students, who made probability estimates for problem situations which had been specifically designed for this study. The main experiment was preceded by a pretest using a subset of the problem situations, for which estimates were made by students from Catholic University.

A. METHOD

The materials used in this experiment included a set of three types of problem situations (RFB, NRFB, and almanac). The problem situations

were preceded by instructions explaining the probability estimation task and by practice items. A questionnaire ended the research session. The complete set of materials used in the experiment is reproduced in Appendix B.

1. RFB Problem Situations

The RFB items used in this study were of the "book-bag and poker chip" variety. A general format was devised and rigorously used for all RFB items, with the specifics being systematically changed from problem to problem. The basic RFB format may be seen in Figure 1, which exhibits one of the problem situations used in the experiment.

The subjects were told in each RFB problem statement to assume that there were two large identical bags each filled with a large number of ordinary poker chips. Some of these chips were red and some were blue. Bag "B" contained *predominantly* blue chips in a certain ratio, while bag "R" contained predominantly red chips in the same ratio. Thus, if bag "B" contained 64% blue chips and 36% red chips, bag "R" contained 64% red chips and 36% blue chips.

A fictitious coin toss for selecting a bag made the prior probability of selection 50% for each bag. Sequential one-chip samples, with replacement and remixing after each sample, were then assumed to have been drawn from the chosen bag. The results of the samples were presented to the subjects simultaneously. The subject's task was to estimate the (revised) probability that a specified bag had been chosen, based on the information given in the sample draws. The problem statements emphasized that the initial probability was 50% and that replacement and remixing occurred

Problem Situation #10

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 85% blue chips and 15% red chips
Bag "R" contains 15% blue chips and 85% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "B" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is blue. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is red. The chip is returned to the bag, the contents are remixed and a fourth chip is drawn. This chip is blue. The chip is returned to the bag, the contents are remixed and a fifth chip is drawn. This chip is blue.

Question: What would you estimate to be the probability that the selected bag is bag "B"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is blue
(3) Third chip is red
(4) Fourth chip is blue
(5) Fifth chip is blue

Answer:____________%

FIGURE 1: An RFB Problem Situation Used in the Experiment

after each draw.

Following the full narrative description given in each problem situation, including the bag color ratios and the draw sequence, a simplified summary of the problem conditions was provided. The subject then used a percentage between 0% and 100% to estimate the probability that a specified bag had been chosen.

There were no instructions given for interpreting the results of the sample draws, except that a note provided after the RFB practice item explained that "since more blue chips were drawn from the bag, it is reasonable to estimate that the probability that bag "B" was chosen is greater than 50% because that bag contains more blue chips than bag "R" does."

For the class of RFB problem situations used in this study, posterior probability is a function of the arithmetical difference between the number of red and blue chips.[1] It was therefore found efficient in constructing the RFB items to use a computer program to generate candidate problem situations as a function of bag content ratio and the excess of one color over the other for all draws within the sample. Consideration was limited to color excesses of one, two or three chips over all draws within a problem. True probabilities for the candidate situations were arranged in sequence, and twenty items were selected to include two true probabilities within each decile. Draw sequences were established so that each sample contained from two to five chips, and the sequence of colors of the draws was randomized. The target bag was then designated for each item, and the overall sequence of the RFB problems was randomized.

---

[1] See Shanteau (1972), for example.

Ten of the RFB problems thus constructed were used in the pretest. Four additional problems included in the main experiment were randomly inserted in the sequence of problems in the experiment. The four added items were selected to have true probabilities in the top and bottom two deciles of the range in order to provide better resolution for extreme probabilities. The basic characteristics of the fourteen RFB items used in the main experiment are summarized in Figure 2. The pretest sequence numbers are shown in the right-hand column.

2. NRFB Problem Situations

The NRFB problems used in the experiment were general information situations constructed so as to have a format similar to the RFB items. The main difference between the RFB and NRFB problems was that the latter lacked a relative frequency base, and thus no true probability could be calculated. In addition, whereas all the data elements in the RFB items were the same (i.e., chip draws), a variety of data elements were used in the NRFB items. An example of a NRFB problem is given in Figure 3.

The nine NRFB problems, which were used in both the pretest and main experiment, included a variety of topics. The varied subject matter of these problem situations may be seen in Appendix B.

It was necessary to devise NRFB situations which fell within the range of experience of the subjects and which were relatively easy for the average person to grasp. In each NRFB problem there were from three to five data elements. To the greatest extent possible, each was constructed to be conditionally independent of the others. These data elements were

ITEM NO.	BAG COMPOSITION (Percentage)				TARGET BAG	DRAW OUTCOMES (In Sequence)					TRUE PROBABILITY	PRETEST ITEM NO.
	BAG "B"		BAG "R"									
	blue	red	blue	red		1	2	3	4	5		
1	66	34	34	66	"R"	red	red	-	-	-	.790	1
2	75	25	25	75	"R"	blue	blue	blue	red	blue	.036	2
3	70	30	30	70	"R"	red	blue	blue	blue	-	.155	-
4	52	48	48	52	"R"	blue	blue	-	-	-	.460	3
5	78	22	22	78	"B"	blue	red	red	-	-	.220	4
6	70	30	30	70	"B"	blue	blue	red	blue	blue	.927	-
7	67	33	33	67	"B"	red	red	red	-	-	.107	5
8	54	46	46	54	"R"	red	red	blue	-	-	.540	6
9	89	11	11	89	"B"	red	red	-	-	-	.015	-
10	85	15	15	85	"B"	red	blue	red	blue	blue	.850	7
11	90	10	10	90	"R"	red	red	blue	red	-	.988	8
12	59	41	41	59	"B"	blue	red	red	red	-	.326	9
13	65	35	35	65	"R"	red	red	red	-	-	.865	-
14	67	33	33	67	"R"	red	red	blue	-	-	.670	10

FIGURE 2: Characteristics of the RFB Problem Situations

Problem Situation #15

Assume that GM sells the same number of 1978 trucks that are sold by non-GM manufacturers. Thus, if in a conversation with a stranger, you found out that he had purchased a new 1978 truck, it would be reasonable to estimate that the probability is 50% that his truck was a GM.

Suppose that, at the beginning of the next model year, GM sustains a massive strike for three months during which no 1979 trucks are produced, while all non-GM truck manufacturers avoid strikes. Further suppose that, midway through the model year, a Ralph Nader-type group issues a report that 1979 GM trucks have serious safety defects. In addition, assume that although 1979 GM trucks increased in price by an average of $190 over the previous year's price, non-GM trucks increased in price by an average of $345.

Sometime in 1980 you meet another stranger whom you find has purchased a new 1979 truck.

Question: What would you estimate to be the probability that it was a GM truck?

Summary

Initial probability estimate: 50%

Information:

(1) GM strike
(2) GM safety questioned
(3) Smaller GM price increase

Answer: ______________________________%

FIGURE 3: An NRFB Problem Situation Used in the Experiment

designed to influence a respondent's probability estimate in a specific direction and within a given range of intensity. The subjects were, for *each problem*, to combine the overall magnitude and direction of the data elements into a single weighted assessment of the posterior probability.

NRFB items can quite readily be constructed to elicit a fairly consistent response by generating only likely (or only unlikely) events conditional on a specific hypothesis so that the overall effect, by almost anyone's inherent weighting scheme, strongly favors one hypothesis. It is more difficult to devise a set of general information NRFB problem situations such that the resultant probability estimates span a range of values. One obvious way to obtain an estimated probability within the middle of the probability range is to construct a set of data elements with a consistent trend, but with low to moderate intensity so that the combination does not drive the response to the extreme. Another way is to devise a carefully balanced set of elements, some of which offset one another. Of course, subjects may differ in their perceptions of individual data elements, in their weighting scheme, or in both.

In an attempt to obtain coverage of the probability spectrum, the experimenters made independent judgments of the general position of the posterior probability for each NRFB problem. Where necessary, adjustments were made to problems in order to provide

(1) agreement in the experimenters' rough assessments

and (2) coverage of the entire (0,1) probability range.

These assessments were partially confirmed when the subjects' average estimated probabilities were calculated for the pretest and the main

experiment. When these were compared with the experimenters' assessments, there was agreement for seven of the nine NRFB items. Of course, it was not expected that there would be complete agreement between the subjects and the experimenters.

3. Almanac Problem Situations

A third type of problem situation employed in this experiment was based on information obtained from the 1977 World Almanac. The almanac items, which could not be put into a format similar to the RFB and NRFB problem situations, were included in the experiment in order to gain information about the feasibility of using RFB-based calibration in problem situations which did not have a structure parallel to that of the RFB problems. Figure 4 provides an example of an almanac problem used in the experiment. A total of eight almanac problems were employed in both the pretest and the main experiment. These problems are reproduced in Appendix B.

Items which could be cast in probabilistic terms were chosen from the almanac. This restricted choice to situations in which a population was broken into constituent elements, such as the population of the U. S. broken down by states. An additional restriction was that a reasonable sounding probability question had to be possible from the information. It seemed fairly reasonable to ask subjects to estimate the probability that a baby born in the U. S. during 1974 was born east of the Mississippi, since the true probability could be established in terms of the relative occurrence of births by states. It would be more than a bit strained, however, to ask the probability that a square mile of U. S. land chosen at random is

Problem Situation #28

FBI figures for 1975 report that there were 116,000 arrests in the U.S. that year for robbery.

Question: What would you estimate to be the probability that a 1975 U. S. arrest for robbery was a man rather than a woman?

Answer: ____________________%

FIGURE 4: An Almanac Problem Situation Used in the Experiment

in Alaska, even though the almanac indicates that 16% of the U. S. landmass is in that state.

B. THE PRETEST

A pretest had the purpose of trying out and evaluating all aspects of the test administration and data analysis procedures. However, the two most important objectives were to determine how much time should be allowed for completion of the problem situations, and to judge the reasonableness of the RFB and NRFB items based on comments solicited from the participants.

Subjects for the pretest were students in an introductory psychology class at Catholic University. Arrangements were made with the instructor to administer the test during a regular class period, which lasted less than an hour. Class members had been informed earlier in the semester that a class period (date unspecified) would be devoted to an experiment in which they would participate. Twenty-two persons took part, but since three were unable to complete all of the items in the time available, only the data from nineteen subjects were included in the analysis.

Instructions, including a brief outline of the probability concepts involved in the experiment, were read aloud to the pretest subjects. The experimental materials were then distributed, and the subjects proceeded to work the three practice items. They were given an opportunity to ask questions before continuing with the main part of the experiment.

The pretest consisted of ten RFB items, nine NRFB items and eight almanac items. A questionnaire comprised the last two pages of the experimental materials. The first page of questions asked the subjects to evaluate the test items for validity and realism. They were requested to evaluate

the bag and chip problems as a group, the NRFB problems individually, and the almanac items as a group. The responses received are summarized in Figure 5. (As can be seen from this figure, not all of the participants completed the questionnaire.) The consensus of the respondents seemed to be that the problems were fairly realistic and the initial 50% probability assumption was not difficult to accept. However, there was considerable uncertainty expressed by the subjects concerning the accuracy of their estimates.

The second questionnaire page asked for background information about the subject's age, sex, major, minor, grade point average, credits in probability, statistics or decision theory and years of education. Subjects were also asked if they thought there was a way to calculate the bag and chip problem probabilities. A place for comments was provided. The questionnaire is included in Appendix B with the other experimental materials.

C. THE MAIN EXPERIMENT

Because the responses of the pretest subjects indicated that the RFB, NRFB and almanac problem situations were reasonably realistic, these items were judged suitable for inclusion in the main experiment. Before subjects were recruited for the experiment, however, an examination of statistical power was made. (See Appendix A.)

Although the existing ten RFB items satisfied the minimum power requirements, four additional RFB problem situations were included in the main experiment. These four items, which were selected to have true probabilities near the two ends of the (0,1) range, were included primarily to have a moderating influence on the number of predicted probability estimates that

SUPPLEMENTARY INFORMATION

1. Some of the problems you have just completed established an initial 50% probability estimate which you were asked to revise on the basis of additional information. Please indicate below how valid you think the initial 50% probability estimate was in each case. Then indicate how sure you feel about the probability estimates you made. Finally, indicate how realistic you feel each situation was. **Place check marks to indicate your evaluation.**

PROBLEM SITUATION	How Was Initial 50% Probability Estimate?			How Sure Were You of Your Estimates?			How Realistic Were Situations?		
	Seemed Okay	Off Slightly	Seriously Off	Quite Accurate	Only Fairly Accurate	Sheer Guess	Fairly Accurate	A Bit Contrived	Completely Hokey
Bag and Chip Problems	16	5		4	16	1	15	6	
GM Trucks	12	7	2	3	14	3	16	4	
Elective Course	14	5	2	9	9	2	16	5	
Street Intersection	17	3	1	10	8	2	16	4	
Illinois Resident	14	6	1	3	14	3	10	9	2
New York Harbor	14	7		3	13	4	15	4	1
Pepsi vs. Coke	17	4		13	5	2	18	3	
Presidential Election	16	4	1	9	9	2	19	1	1
Mrs. Jones' Baby	12	5	4	8	10	3	9	7	5
Elevator	11	8	2	4	11	4	11	6	2
Problems with No Initial Probability Estimate				1	15	2	13	5	

FIGURE 5: Summary of Questionnaire Responses Received from the Pretest Subjects

were outside this interval. (A number of such out-of-range predictions were noted in analysis of the pretest data.)

Students from The Pennsylvania State University were recruited as subjects in the experiment by means of an advertisement in the student newspaper. This ad promised ten dollars for participation in a two-hour "...decision-making study involving a simple pencil and paper task." To insure getting at least 25 usable subjects, allowing for various types of attrition, a total of 51 respondents were recruited. Surprisingly, only four of them failed to appear at the session, so a total of 47 subjects took part.

The experiment was administered in a single group session with subjects separated from each other at large tables. The session was continuously monitored by two experimenters.

After all subjects were seated, instructions were presented orally. Then, the printed instructions and practice items (Appendix B) were distributed. Each subject read the instructions and worked on the practice items independently.

To insure understanding of the procedure, the group was asked for a show of hands for each of the practice items (e.g., "Raise your hand if your estimate for Example A was greater than 50%.") For Example C (an RFB problem) no one reported an estimate less than 50% (i.e., a posterior probability in the wrong direction). Had there been, further instructions would have been given.

Subjects were instructed to work the problems in sequence, to check their work when finished, to make sure all items were answered, to fill out the questionnaire at the end, and to wait until everyone had finished.

The experimental materials, which were stapled in the upper left hand corner, were then distributed and the experiment began. The sets of materials were identified only by a serial number on every page to insure that pages could be collated should any set come apart. Anonimity was thus preserved, an assurance the subjects had been given during the instruction period.

All subjects finished within the two hours allotted, with some early finishers having to wait fifteen mmnutes or more. All experimental materials were then collected and each subject was paid for participating.

IV. DATA ANALYSIS

The fundamental assumption on which this study was based is that it should be possible to construct (for each subject) a predictive equation, derived from estimates made for RFB problem situations, by which assessments of NRFB probabilities may be calibrated. The statistical procedure for testing this assumption is outlined in Appendix A. Although the probability assessments of the subjects in both the pretest and the main experiment were subjected to statistical analysis, only the results of the main experiment are discussed because of the preliminary nature of the pretest.

As the first step in data analysis, the probability assessments were subjected to a preliminary examination aimed at uncovering any anomalies. After this examination was completed, regression techniques were used to obtain calibration equations for predicting true probabilities from estimated probabilities. These equations, in conjunction with the experimental data, were then used to test the study assumption by means of analysis of variance (ANOVA). The following sections describe these three phases of data analysis.

A. PRELIMINARY EXAMINATION OF THE DATA

Before any predictive (i.e., calibration) equations were estimated, the experimental data was subjected to a preliminary examination. This examination, which was based to a large extent on various data plots, was conducted primarily to check for any anomalies within the probability

assessments. Immediately obvious was one subject who had estimated 50% as the probability for every one of the problem situations. Because of this, the subject's responses were discarded, so that subsequent data analysis was based on a total of 46 subjects.

The data examination also revealed that there were extensive individual differences in the subjects' assessments. In many instances probability estimates were found to be at opposite ends of the probability continuum for the same item. This was not in itself surprising, but while many of the discrepant responses may well have been due to real and consistent differences among subjects' probability perceptions, they may also have been the result of carelessness in responding. This possibility is discussed in more detail in the following sections.

1. RFB Reversals

In both the pretest and the main experiment the subjects were cautioned to pay particular attention to the conditions specified in each RFB problem situation. This seemed advisable because although the bag and chip problems were all of the same format, the composition of the bags varied from item to item, as did the identity of the bag which was to be considered in making the probability estimate response. If a problem asked for the probability of bag "B" but the subject were to make an estimate instead with respect to bag "R", a serious distortion would result. Such responses may be referred to as reversals. An RFB reversal response may be defined by comparing the subject's estimate with the true probability. Specifically, a reversal is either a response which is greater than 50% when the true probability is less than 50%, or a response which is less than 50% when the

true probability is greater than 50%.

Reversals were only a minor factor in the pretest. Fourteen of the nineteen pretest subjects were found to have no RFB reversals whatsoever, while for three of the five who did have reversals, the effect was insignificant. At the time it was thought that the reversals might have resulted from time pressure, and that with ample time the number of reversals would be minimized. To guard against reversals the procedure in the main experiment was revised to emphasize in the instructions that the conditions for each RFB item should be considered carefully. A poll of the subjects after completing the practice items did not indicate any apparent need to further reinforce the instructions.

However, after the data from the main experiment was examined, it was apparent that reversals could cause a serious analysis problem. Although only 6% of the 644 total RFB estimates were reversals, almost half of the participants made at least one reversal response. Because reversals could potentially have a significant effect on the outcome of the experiment, the resulting data was subsequently analyzed two ways. In one analysis the responses were taken as given, while in the other analysis each RFB reversal was "revised" by subtracting the estimated probability from unity.

The effects of a reversal may be seen by examining Figure 6, which presents a plot of true probability against estimated probability for one of the subjects in the main experiment. As can be seen, the point (.88, .015) seriously distorts what is otherwise a reasonable linear relationship. Revision of the reversal corrects this distortion.

As an explanation of the reversals, a fatigue effect might be postulated. However, although there were slightly more reversals among the

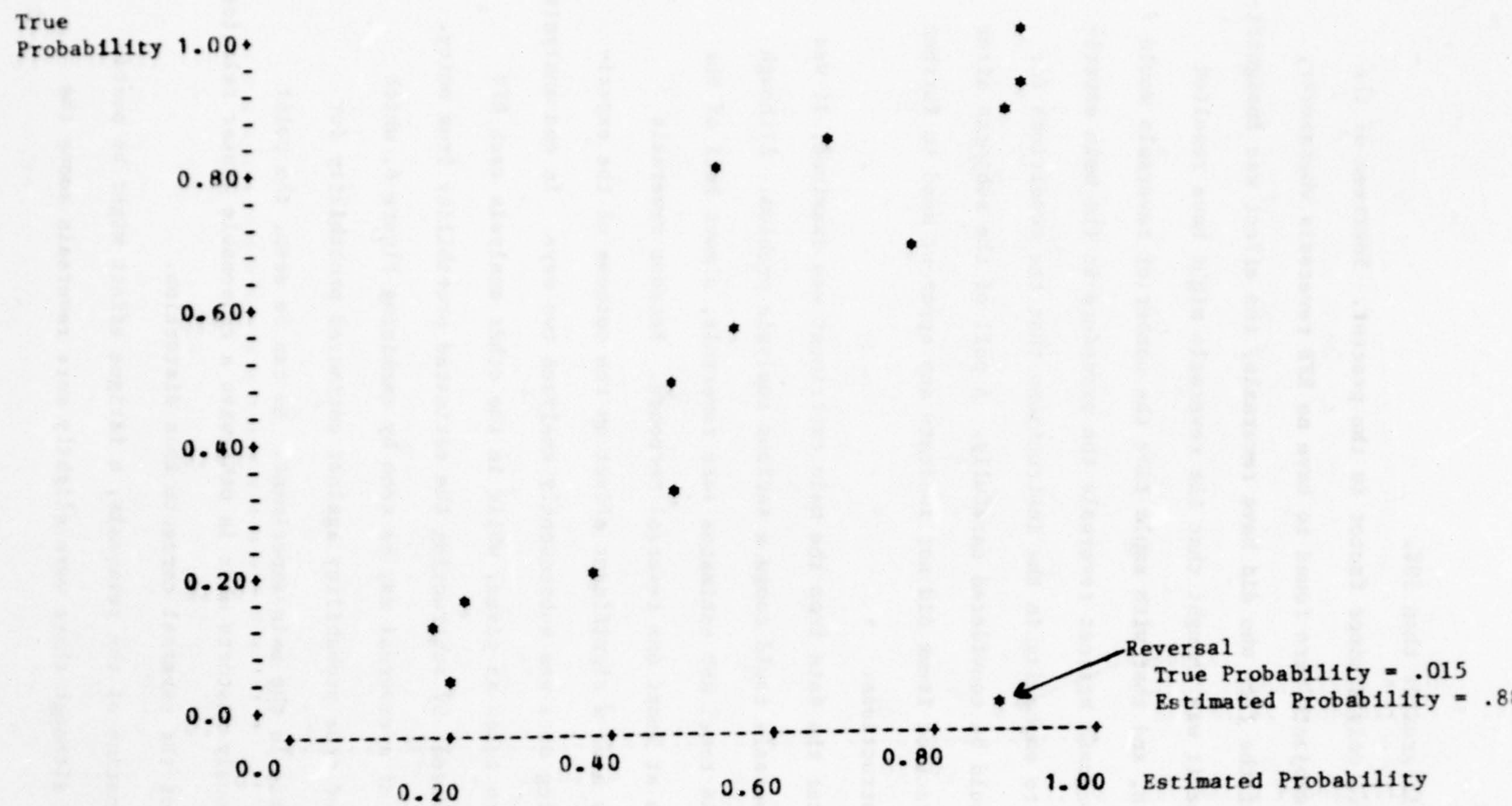


FIGURE 6: An Example of an RFB Reversal

last seven RFB items than among the first seven, there were more among the first four items than among the last four. It is interesting to note that there were 23 reversals of items which had true probability less than 50% and only 13 reversals of items in which the true probability was greater than 50%, perhaps indicating a tendency to estimate the probability of the most likely hypothesis. Nonetheless, these results are not statistically significant. Since the reversal patterns cannot be satisfactorily explained by other assumptions, it may be that lack of attention and carelessness were the causes.

2. NRFB Reversals

Just as it is possible for a subject to give a response to an RFB item which is in a sense the reverse of that subject's underlying estimate, it is conceivable that a response to an NRFB item might also be a "reversal." In other words, the subject might be estimating the probability of the complementary hypothesis. However, there is no "true" probability against which to compare NRFB estimates as there is for the RFB book-bag and poker chip problems used in this study. Although it is possible to use the arithmetic mean over all subjects as the criterion for defining an NRFB reversal, the justification is much more tenuous. Nevertheless, if NRFB reversals were defined in this way, a considerable number of NRFB estimates could be classified as reversals.

B. ESTIMATION OF PREDICTIVE EQUATIONS

In the investigation of the functional (i.e., predictive) relationship

between true and estimated probability, attention was restricted to relationships which are intrinsically linear [Draper and Smith (1966)] and, more specifically, which are or may be transformed to a low-order polynomial. Thus, the added complexities of nonlinear estimation were avoided. This restriction did not impose extreme limitations; it permitted consideration of the dependent variable as any polynomial function of the independent variable and allowed the use of any chosen transformations of the true and estimated probabilities.

In the analysis of the data, attention was concentrated on probability and log odds variables. Both p (true probability) and $\ln[p/(1 - p)]$ were plotted, as dependent variables, against r (estimated probability) and $\ln[r/(1 - r)]$ as independent variables. Linear, quadratic and cubic regression equations were fitted for all four combinations of these dependent and independent variables for each subject.

The best functional realtionships were obtained using true probability versus estimated probability or true log odds versus estimated log odds. Mixing probability with log odds provided less useful calibration equations, as judged by observed F and R^2 values. In general, linear equations were found to provide a reasonable fit, since linear terms were usually highly significant while quadratic and cubic terms were nonsignificant. However, the predictive ability (measured by R^2 and residual mean square) of the regression equation was disappointing for those subjects who had RFB reversals. When these reversals were revised, major improvement was noted. For example, for probabilities as variables, the minimum R^2 value was .755 . Figure 7 lists the observed F and R^2 values for the linear regression equations corresponding to each of the 46 subjects based on the revised

Subject	Probability vs. Probability		Log Odds vs. Log Odds		Subject	Probability vs. Probability		Log Odds vs. Log Odds	
	F	R^2	F	R^2		F	R^2	F	R^2
1*	49.2	.788	46.0	.821	24*	91.9	.888	68.4	.865
2*	146.1	.926	68.0	.865	25*	137.3	.895	45.8	.790
3*	56.8	.808	71.4	.823	26*	118.5	.766	64.3	.704
4*	79.9	.866	42.6	.635	27*	771.1	.988	124.9	.924
5*	134.0	.784	71.8	.694	28	459.6	.967	157.4	.932
6	111.4	.905	139.6	.927	29*	47.8	.764	22.6	.566
7*	68.0	.812	35.8	.755	30	114.1	.864	61.1	.620
8*	79.6	.852	69.4	.866	31*	167.6	.915	113.7	.898
9*	154.3	.886	39.0	.710	32*	462.6	.940	255.6	.945
10	92.4	.876	65.1	.862	33*	111.2	.876	122.3	.916
11*	77.2	.883	31.2	.718	34	196.5	.947	107.8	.910
12	65.0	.832	26.4	.575	35	100.0	.898	81.5	.872
13	121.2	.895	57.7	.852	36	91.4	.895	54.0	.827
14	49.7	.833	39.4	.794	37	319.2	.965	89.6	.885
15*	164.3	.897	133.0	.919	38	185.3	.901	60.9	.781
16*	124.5	.920	44.6	.807	39	153.7	.898	211.9	.950
17	366.1	.971	90.0	.899	40	68.0	.858	59.4	.847
18	99.3	.898	60.8	.856	41	167.5	.981	41.9	.769
19*	82.5	.877	44.8	.797	42	102.7	.879	46.8	.781
20	151.6	.931	42.0	.806	43*	95.9	.809	56.7	.610
21*	44.8	.755	29.7	.635	44	144.2	.914	85.9	.891
22*	164.4	.919	81.6	.883	45	77.4	.848	32.9	.717
23*	55.5	.782	25.3	.666	46	42.8	.764	30.2	.746

* Denotes Subjects without any RFB Reversals

FIGURE 7: Observed F and R^2 Values for Linear Regression Equation (Revised RFB Data)

RFB data.

As expected, there could be instances when a predicted probability computed from a regression equation would fall outside the (0,1) range over which probabilities are defined. To lessen the effect of out-of-range predicted probabilities, these probabilities could, of course, be constrained to fall in the (0,1) range by setting values larger than one equal to one and values less than zero equal to zero.

In view of the previous discussion, three calibration equations for probability prediction were considered for each subject:

(1) an equation based on the original RFB estimates,

(2) an equation based on the revised (reversal-corrected) RFB estimates,

and (3) an equation based on revised RFB estimates with predicted values subject to range constraints.

In addition, two equations for log odds prediction based on a conversion of the probability estimates to log odds were considered: one derived from the original unrevised RFB estimates, and the other based on revision of RFB reversals.

By using R^2 as the criterion, it was determined that some subjects' RFB estimates were consistently much more highly correlated with the true probability than others. There were fourteen subjects for whom the R^2 value based on revised RFB estimates was .90 or larger. Calibration equations were therefore obtained for this group of "most consistent estimators" as well as for the full complement of 46 main experiment subjects and for the subset of 24 subjects having no RFB reversals.

C. EXPERIMENTAL RESULTS

Probability (and log odds) predictions were based, of course, on the linear regression equations derived from estimated probabilities in RFB problem situations. Predictions were made for the primary set of interest (the nine NRFB items) and for the eight almanac problems. As a means of validating the fundamental concept of this study, predictions were also made using linear regression equations derived from the first ten of the RFB items (an arbitrary selection) to predict values for the remaining four RFB items. The resulting predictions in each of these three cases were subjected to the hypothesis tests outlined in Appendix A.

1. Predictions of NRFB Probabilities

To examine the assumption that assessments of NRFB probabilities may be calibrated by means of an RFB-based predictive equation, the experimental results were statistically evaluated by ANOVA for each individual NRFB problem across all subjects and in an overall evaluation across all NRFB problems and all subjects. These individual and overall analyses were also performed for a subset of the 24 subjects without RFB reversals and for the 14 most consistent estimators. In addition, three different data sets were considered. The first was based on the subjects' estimates as given, while in the second the RFB reversals were revised. The third was based on revision of both RFB reversals and NRFB reversals.

Figure 8 shows the results of the overall statistical tests. Large F-ratios and low probabilities, normally considered desirable as a research

PREDICTION	SOURCE OF VARIATION	ALL 46 SUBJECTS: All Original Estimates	ALL 46 SUBJECTS: RFB Estimates Revised	ALL 46 SUBJECTS: All Estimates Revised	24 SUBJECTS WITHOUT RFB REVERSALS: All Original Estimates	24 SUBJECTS WITHOUT RFB REVERSALS: NRFB Estimates Revised	14 MOST CONSISTENT PROBABILITY ESTIMATORS: RFB Estimates Revised	14 MOST CONSISTENT PROBABILITY ESTIMATORS: All Estimates Revised
PROBABILITY (UNCONSTRAINED)	BETWEEN SUBJECTS	2.70 (.000)	3.83 (.000)	2.16 (.000)	3.66 (.000)	2.80 (.000)	3.37 (.000)	2.75 (.001)
	SUBJECTS X PROBLEMS	2.72 (.000)	5.10 (.000)	3.50 (.000)	5.81 (.000)	4.67 (.000)	5.16 (.000)	2.55 (.000)
LOG ODDS	BETWEEN SUBJECTS	3.11 (.000)	3.88 (.000)	2.80 (.000)	4.59 (.000)	3.67 (.000)	2.64 (.002)	2.11 (.016)
	SUBJECTS X PROBLEMS	3.18 (.000)	4.59 (.000)	3.75 (.000)	5.76 (.000)	5.20 (.000)	3.14 (.000)	2.05 (.000)
PROBABILITY (CONSTRAINED)	BETWEEN SUBJECTS	1.88 (.001)	2.60 (.000)	1.26 (.125)	2.14 (.002)	1.52 (.063)	2.90 (.001)	2.50 (.004)
	SUBJECTS X PROBLEMS	1.30 (.003)	2.54 (.000)	1.17 (.049)	1.94 (.000)	1.07 (.302)	4.69 (.000)	2.06 (.000)

FIGURE 8: Observed F-ratios for Overall Test on NRFB Problems
(Corresponding p-values in Parentheses)

result, here imply significantly large differences between predicted probabilities (or log odds) for different subjects, which tends to reject the assumption that the RFB-based models would result in predictions which were all unbiased estimates of the same, unknown, probability for any given problem. Thus, in the current situation, small F values and high probabilities support the underlying assumption.

From an examination of the figure, it can be seen that overall agreement between predictions of the NRFB probabilities was, in general, unsatisfactory. With all reversals revised and range-constrained probabilities used, the underlying assumption would not have been rejected for the set of all 46 subjects and for the subset of 24 subjects without RFB reversals. However, these results were based on the revision of NRFB reversals, which seems to be an extremely tenuous procedure.

2. Prediction of Almanac Probabilities

Almanac items of the type employed in this experiment were of some interest because, although they did not have a general information format similar to the RFB and NRFB problems, they did have true probabilities associated with them. Thus, the items were included in this experiment to provide supplementary information.

Subjects' estimates correlated poorly with the true probabilities for most of the almanac items. More importantly, as Figure 9 indicates, RFB-derived predictive models provided an unstable basis for calibrating probability assessments in the almanac situations. This result was expected, however, because of the dissimilarity between the almanac and RFB format. While the RFB situations were designed as two-hypothesis

PREDICTION	SOURCE OF VARIATION	ALL 46 SUBJECTS	24 SUBJECTS WITHOUT RFB REVERSALS	14 MOST CONSISTENT PROBABILITY ESTIMATORS
PROBABILITY (UNCONSTRAINED)	BETWEEN SUBJECTS	9.85 (.000)	6.14 (.000)	15.48 (.000)
	SUBJECTS X PROBLEMS	4.31 (.000)	4.98 (.000)	4.93 (.000)
LOG ODDS	BETWEEN SUBJECTS	9.83 (.000)	7.27 (.000)	10.92 (.000)
	SUBJECTS X PROBLEMS	4.30 (.000)	4.91 (.000)	3.75 (.000)
PROBABILITY (CONSTRAINED)	BETWEEN SUBJECTS	6.70 (.000)	3.61 (.000)	13.24 (.000)
	SUBJECTS X PROBLEMS	2.34 (.000)	2.19 (.000)	4.38 (.000)

FIGURE 9: Observed F-ratios for Overall Test on Almanac Problems (Corresponding p-values in Parentheses)

problems where a prior probability of 50% was to be revised in the light of new data, the almanac problems asked for a direct probability assessment.

3. Prediction of RFB Probabilities

Although the failure of the RFB-based model when applied to the almanac items was not surprising, their poor performance on the NRFB problems cast considerable doubt on the overall validity of the suggested calibration approach. Of course, success or failure in predicting NRFB probabilities using models derived from RFB probability estimates may depend as much on the similarity of the two estimation tasks as it does on the development of a satisfactory calibration equation.

In order to gain insight into this matter, additional analyses were performed on the reversal-corrected RFB estimates obtained in the experiment. The first ten RFB items were used to derive equations of the subjects' estimation behavior, and these were then used in adjusting estimates made on the last four RFB items to arrive at predicted values. Therefore, the data used for constructing the equations was independent of that used for testing them. Since the task was identical for each group of items, the only variable left to explain failure would be an inadequate calibration equation.

As Figure 10 shows, good results were obtained for probability models and for log odds models. The best results were obtained when predicted probabilities were constrained to the (0,1) range. In view of the almost complete lack of success with the NRFB predictions, these results were unexpected.

PREDICTION	SOURCE OF VARIATION	ALL 46 SUBJECTS	24 SUBJECTS WITHOUT RFB REVERSALS	14 MOST CONSISTENT PROBABILITY ESTIMATORS
PROBABILITY (UNCONSTRAINED)	BETWEEN SUBJECTS	1.27 (.123)	1.34 (.146)	.82 (.638)
	SUBJECTS X PROBLEMS	.76 (.969)	.84 (.798)	.70 (.898)
LOG ODDS	BETWEEN SUBJECTS	1.45 (.036)	1.76 (.021)	.57 (.873)
	SUBJECTS X PROBLEMS	1.06 (.333)	1.25 (.121)	.78 (.810)
PROBABILITY (CONSTRAINED)	BETWEEN SUBJECTS	.92 (.622)	.84 (.678)	.64 (.816)
	SUBJECTS X PROBLEMS	.52 (.999)	.55 (.998)	.58 (.973)

FIGURE 10: Observed F-ratios for Overall Test on Four RFB Problems (Corresponding p-values in Parentheses)

It should be noted that the tests indicate relatively small variability between the set of predictions obtained for a given problem by means of the RFB-derived models. Thus, they strongly support the proposition that for any given RFB problem the models result in unbiased estimates of the same population parameter.

However, the tests provide no comparison of predictions with true values. Figure 11 graphically indicates (for the probability models) the average estimates, the average predictions, and the true probability values for each of the four RFB problems. As can be seen, in all except the last problem the predictions obtained from application of the RFB-based regression equations were much closer to the true values than were the estimates obtained directly from the subjects.

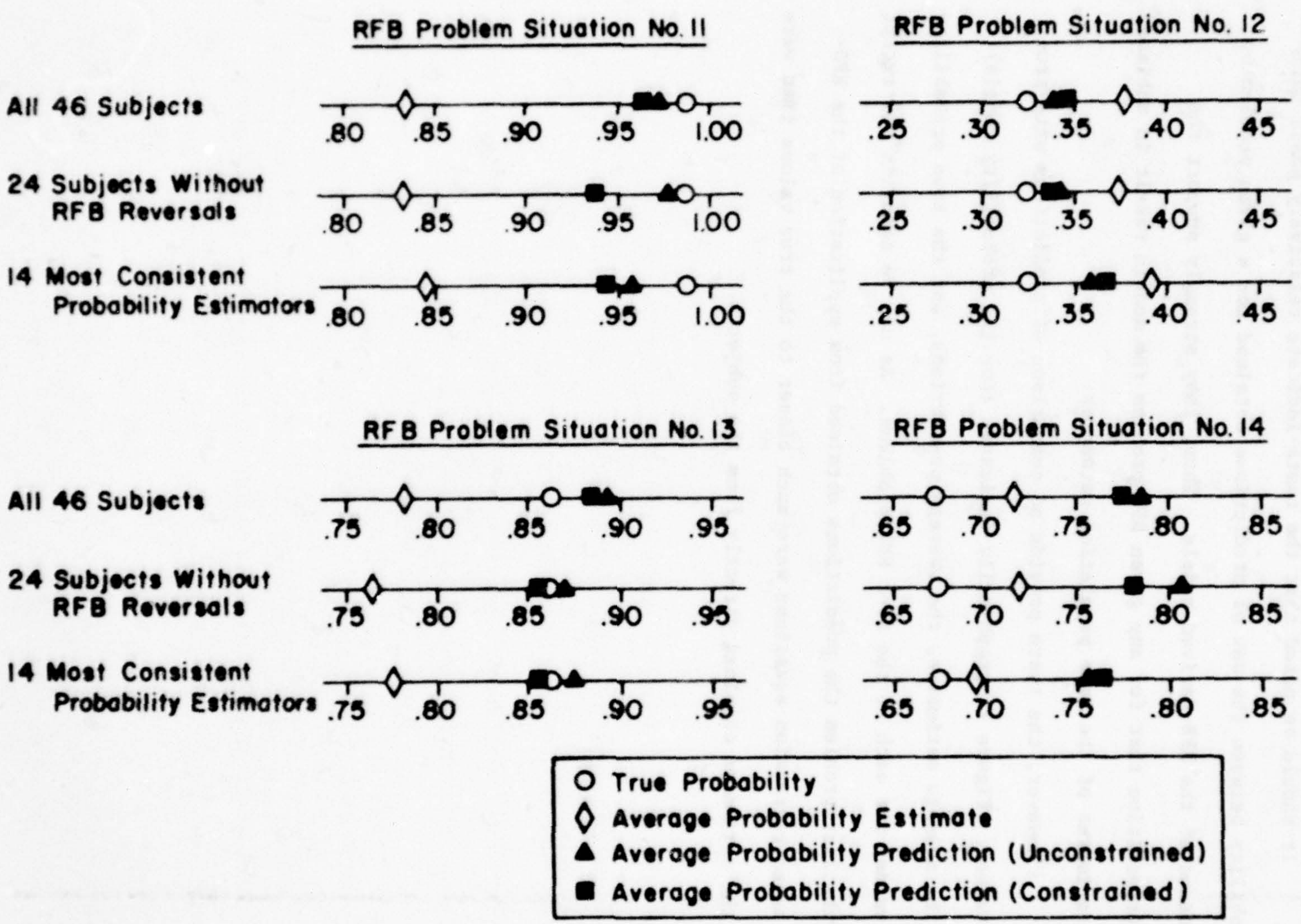


FIGURE 11: A Comparison of True Probabilities, Estimates and Predictions for the Four RFB Problems

V. SUMMARY AND DISCUSSION

The experiment conducted in this study had as its main objective the investigation of the hypothesis that probability estimates made for a set of "book-bag and poker chip" relative frequency base (RFB) tasks could be used to derive satisfactory models for providing accurate probability assessments for a set of general information non-relative frequency base (NRFB) problem situations. A fairly comprehensive set of predictive models was employed, some of which proved to be unsatisfactory and were eliminated from further consideration. In fact, it was never possible to make satisfactory predictions of NRFB probabilities when the subjects' raw RFB probability estimates were used in constructing a model.

Some improvement was obtained by revising all RFB estimation reversals and constraining the range of all predicted probabilities. However, it was not until NRFB reversals were also revised that any satisfactory results were obtained. Unfortunately, any proposed justification for this last step is questionable.

Despite this lack of success, there is strong evidence that the underlying approach may be valid if the probability estimation tasks to which a predictive model is applied are identical to those used in model construction. This evidence is provided by the results of using estimates made on the first ten of the RFB problems to construct models which were then used to make predictions for the remaining four RFB items. The predictions thus made were found to support the suggested approach.

It should be recalled that the RFB and NRFB problem situations were

specifically designed to be of parallel construction. Of necessity, however, the data and hypotheses in the RFB problems differed from those in the NRFB problems. If the subjects perceived the NRFB problems as inherently distinct from the RFB problems, this might account for the failure of RFB-based models to provide adequate predictions of NRFB probabilities.

Another explanation for the experimental results might be that in real-world situations (such as the NRFB problems purport to be) probabilities are assessed on the basis of "common sense" and experience in dealing with past problems. Each individual brings to each problem a different experiential framework (the result of many events and their interactions) which causes him to approach the problem from his own perspective. Each individual's experiential framework will affect the cognitive processes as he refers to the framework in estimating a probability. The situation is further complicated if the subject perceived each NRFB problem as being so different from every other NRFB problem that it was necessary for him to approach each task from a different viewpoint within his experiential framework.

On the other hand, in the case of the "book-bag and poker chip" RFB problems, more similarity of experience would be expected on a problem-to-problem basis. Furthermore, a uniformity of approach to these problems was somewhat guaranteed by the fact that they were of identical structure. Therefore, even without a well-defined mathematical procedure to follow, it would be expected that an individual, once having formed an empirical procedure of his own geared to the particular problem-solving task, would solve all the identically-structured RFB problems according to that procedure.

In summary, it is reasonable to conclude from this experiment that models derived from RFB estimates can provide satisfactory adjustments of probability estimates made for another set of tasks, if these tasks have the identical structure of the RFB problem. It should be noted that although identical structure has been shown to be a sufficient condition for successful performance of the suggested approach, this does not mean that it is a necessary condition. If it were, the approach would, of course, be of severely limited applicability in real-world situations, since few decision-making tasks have a "book-bag and poker chip" format. However, it would be unwise to discount the feasibility of developing a practical RFB-based model for adjusting real-world probability estimates until the approach were tested with participants and problems in a field of common expertise (such as intelligence analysis or weather forecasting, for example) where more uniformity of background and experience exists.

VI. REFERENCES

Beach, L. R., and Wise, J. A. (1969). "Subjective Probability Revision and Subsequent Decisions", Journal of Experimental Psychology, Vol. 81, pp. 561-565.

Brown, T. A., and Shuford, E. H. (1973). "Quantifying Uncertainty into Numerical Probabilities for the Reporting of Intelligence", Rand Corporation, Report No. R-1185-ARPA.

Brown, R. V., Kahr, A. S., and Peterson, C. (1974). Decision Analysis for the Manager, Holt, Rinehart and Winston, New York.

Chinnis, J. O., and Peterson, C. R. (1970). "Nonstationary Processes and Conservative Inference", Journal of Experimental Psychology, Vol. 84, pp. 248-251.

Draper, N. R., and Smith, H. (1966). Applied Regression Analysis, John Wiley and Sons, Inc., New York.

Edwards, W. (1966a). "Introduction", IEEE Transactions on Human Factors in Electronics, Vol. HFE-7, pp. 1-6.

Edwards, W. (1966b). "Nonconservative Probabilistic Information Processing Systems", University of Michigan, Institute of Science and Technology, Report No. ESD-TR-66-404 (AD 647-092).

Goodman, B. C. (1972). "Action Selection and Likelihood Ratio Estimation by Individuals and Groups", Organizational Behavior and Human Performance, Vol. 7, pp. 121-141.

Johnson, E. M. (1973). "Numerical Encoding of Qualitative Expressions of Uncertainty", U. S. Army Research Institute for the Behavioral and Social Sciences, Technical Paper 250 (AD 780-814).

Johnson, E. M. (1974). "The Effect of Data Source Reliability on Intuitive Inference", U. S. Army Research Institute for the Behavioral and Social Sciences, Technical Paper 251 (AD 784-097).

Keeping, E. S. (1962). Introduction to Statistical Inference, Van Nostrand Company, Inc., Princeton.

Peterson, C. R., Schneider, R. J., and Miller, A. J. (1965). "Sample Size and the Revision of Subjective Probabilities", Journal of Experimental Psychology, Vol. 69, pp. 522-527.

Peterson, C. R., DuCharme, W. M., and Edwards, W. (1968). "Sampling Distributions and Probability Revisions", Journal of Experimental Psychology, Vol. 76, pp. 236-243.

Peterson, C. R., and Swenson, R. G. (1968). "Intuitive Statistical Inferences about Diffuse Hypotheses", Organizational Behavior and Human Performance, Vol. 3, pp. 1-11.

Phillips, L. D., Hays, W. L., and Edwards, W. (1966). "Conservatism in Complex Probabilistic Inference", IEEE Transactions on Human Factors in Electronics, Vol. HFE-7, pp. 7-18.

Scheffe, H. (1959). The Analysis of Variance, John Wiley and Sons, Inc., New York.

Shanteau, J. (1972). "Descriptive Versus Normative Models of Sequential Inference Judgment", Journal of Experimental Psychology, Vol. 93, pp. 63-68.

Sibley, W. L. (1974). "A Prototype Computer Program for Interactive Computer-Administered Admissible Probability Measurement", Rand Corporation, Report No. R-1258-ARPA.

Theil, H. (1967). Economics and Information Theory, North-Holland Publishing Co., Amsterdam.

Warner, S. L. (1963). "Multivariate Regression of Dummy Variates under Normality Assumptions", Journal of the American Statistical Association, Vol. 58, pp. 1054-1063.

Wheeler, G., and Beach, L. R. (1968). "Subjective Sampling Distributions and Conservatism", Organizational Behavior and Human Performance, Vol. 3, pp. 36-46.

Winer, B. L. (1971). Statistical Principles in Experimental Design (Second Edition), McGraw-Hill, Inc., New York.

Winkler, R. L., and Murphy, A. H. (1973). "Experiments in the Laboratory and the Real World", Organizational Behavior and Human Performance, Vol. 10, pp. 252-270.

APPENDIX A: STATISTICAL CONSIDERATIONS

As outlined in section II, it was assumed that a functional relationship between true probability and estimated probability can be identified for any given individual. It was also assumed that this relationship could be estimated by means of regression techniques. Because the dependent variable would be true probability (or, perhaps, a transformation), some questions are raised if the overall regression situation is viewed from a classical statistical standpoint.

In a classical statistical framework, the true probability p for any problem is a constant value and therefore does not have a probability distribution (except, of course, a degenerate one). Because the dependent variable under consideration is not, in fact, a random variable, the regression of true probability on estimated probability has no meaning from a classical point of view. [See Keeping (1962), for example.] If, however, the problem is viewed in a Bayesian framework, it can be verified that the true probability p does have a probability distribution.

Consider, for example, the class of problems which, for person i, give rise to the same estimated probability r_0. In general, problems within this class have different true probabilities. Thus, in a Bayesian sense, the true probability of any problem within this class is a random variable. Furthermore, as this argument is repeated for all values of r_0 in the interval (0,1), it is reasonable to assume that the expected value of p is conditional on r_0. Therefore, in this framework, the underlying assumptions of regression are satisfied. Attention may now be turned to

determination of the functional relationship between true probability and estimated probability.

A. FUNCTIONAL RELATIONSHIPS

Given the regression situation with dependent variable p (true probability) and independent variable r (estimated probability), there are two major approaches to postulating the mathematical form of the relationship between the two variables. The first is based on apriori assumptions; the second is based on examination of the scatter plot between the two variables or their transforms. Both approaches have been adopted in this study.

The simplest functional relationship between two variables y and x is, of course, a straight line. The regression model for this relationship would be

$$E(y) = \alpha + \beta x .$$

This is, in essence, the relationship assumed by Sibley (1974) in conjunction with the postulated realism line. However, because a linear relationship does not place restrictions on the predicted values of y , the logistic function has been suggested [Warner (1963), Theil (1967)] as a more appropriate model when probabilities are involved.

The logistic function, which restricts predicted y values to (0,1), takes the form

$$E(y) = \{1 + \exp[-(\alpha + \beta x)]\}^{-1}.$$

If, for example, the observed data in Figure 1 of Beach and Wise (1969)

is examined, it can be seen that the relationship between estimated probability r and true probability p appears to be reasonably well-represented by a fitted logistic function of the form

$$r = \{1 + \exp[-(a + bp)]\}^{-1}. \qquad (1)$$

An "eyeball" examination suggests that values of a = -1.0 and b = 2.0 would provide a reasonably good fit to this data. (Of course, actually fitting a regression equation to the data would provide more accurate values of a and b .)

The relationship between r and p given in (1) may be expressed equivalently (after a few algebraic steps) in the form

$$p = c + d \cdot \ln[r/(1 - r)] \qquad (2)$$

Thus, it is reasonable to investigate the use of p as dependent variable and ln[r/(1 - r)] as independent variable, since it may be that the resulting regression equation in this case is linear.

Other research [Peterson, Schneider and Miller (1969), for example] suggests the relationship

$$\text{BLLR} = k \cdot \text{SLLR} \qquad (3)$$

where:

BLLR denotes Bayesian log likelihood ratio

SLLR denotes subjective log likelihood ratio

and k is a constant.

When the prior probability of each of the two hypotheses under consideration is 50%, BLLR is equivalent to ln[p/(1 - p)] and SLLR is equivalent

to $\ln[r/(1 - r)]$, so that (3) may be written as

$$\ln[p/(1 - p)] = k \cdot \ln[r/(1 - r)] \ . \qquad (4)$$

Thus, this transformation of p and of r should be considered.

Of course, in actually fitting a relationship as indicated in (4), a constant term would be used so that the resulting equation would be of the form

$$\ln[p/(1 - p)] = h + k \cdot \ln[r/(1 - r)] \ . \qquad (5)$$

It should be noted that this equation is algebraicly equivalent to

$$p = \{1 + \exp[-(a + b \cdot \ln[r/(1 - r)])]\}^{-1}$$

which is the form of the suggested logistic function, but in terms of a transformed independent variable.

In view of the foregoing discussion, it appears that attention should be focused on the untransformed variables p and r and on the transformed variables $\ln[p/(1 - p)]$ and $\ln[r/(1 - r)]$. Specifically, the following three combinations of dependent and independent variables are of primary interest in the examination of the functional relationship between estimated and true probability:

Dependent variable	Independent variable
p	r
p	$\ln[r/(1 - r)]$
$\ln[p/(1 - p)]$	$\ln[r/(1 - r)]$

The untransformed variables are, of course, measured on the original

scale over the interval (0,1). The transformed variables are, because of the uniform prior, equivalent to log likelihood ratio and to log odds, which have been used extensively in probability studies [Peterson and Swenson (1968), Peterson, DuCharme, and Edwards (1968), Johnson (1974), for example]. As Edwards (1966a) points out, log likelihood ratio and log odds often tend to be the most useful variables for analysis of probability estimation experiments which involve discrete hypotheses.

B. HYPOTHESIS TESTING

Assume that a total of I subjects participate in the experiment, with each subject providing an estimated probability for each of M relative frequency base (RFB) problems and K nonrelative frequency base (NRFB) problems. Assume that, based on the M RFB problems, a fitted equation $\hat{p} = \hat{f}_i(r)$ is found to provide a good approximation to the relationship between true probability p and estimated probability r for subject i .

As discussed previously, either the true probability p or some transformation of it may be used as the dependent variable. For the remainder of this section, however, the discussion assumes the untransformed probability p to be the dependent variable. This assumption is made only to simplify the discussion and results in no lack of generality, since a similar analysis holds for any transformation of p .

For the kth NRFB problem (where the true probability is, of course, unknown), consider the probabilities $r_{1k}, \ldots, r_{Ik}$ estimated by the I subjects. The functions $\hat{f}_i(r_{ik})$ for $i = 1, \ldots, I$ may be used to provide "corrected" estimates of the true (unknown) probability p_k . If the

use of these functions to obtain corrected estimates is valid, then the estimates $\hat{p}_{1k},\ldots,\hat{p}_{Ik}$ should be unbiased estimates of p_k. In that event, the variability among the $\hat{p}_{ik}$'s would be totally accounted for by the random error inherent in any statistical quantity.

In general, however, each $\hat{p}_{ik}$ is an unbiased estimate of some (unknown) probability p_{ik}. If all the p_{ik}'s are not equal (i.e., the functions do not result in unbiased estimates of the same probability p_k), the the variability among the $\hat{p}_{ik}$'s will, in addition to random error, be attributed to the true differences in the p_{ik}'s. Thus, the "corrected" estimates may be examined in a hypothesis testing framework with the following hypotheses:

$$H_0: \; p_{1k} = \ldots = p_{Ik}$$

$$H_1: \; \text{All the } p_{ik}\text{'s are not equal.}$$

It was noted previously that the fitted function $\hat{f}_i(r)$ provided the estimate $\hat{p}_{ik}$. In addition, the actual fitting of the function provides an estimate $\hat{\sigma}^2_{ik}$ of the underlying variance. This estimate is, of course, given by the residual mean square. If $f_i(r)$ is assumed to be a two-parameter function, $\hat{\sigma}^2_{ik}$ will have M - 2 degrees of freedom.

A test of hypothesis H_0 against hypothesis H_1 may be based on the F-statistic

$$F = MSB/MSE$$

where MSB is the between-subjects mean square

$$MSB = \frac{1}{I-1}\sum_i [\hat{p}_{ik} - (\sum_i \hat{p}_{ik}/I)]^2$$

and MSE is the error mean square formed from pooling the estimates $\hat{\sigma}^2_{ik}$

$$MSE = \frac{1}{I} \sum_{i} \hat{\sigma}^2_{ik} .$$

Assuming that each $\sigma^2_{ik} = \sigma^2$ (a reasonable assumption in view of the robustness of the F-test to inequality of variance), the F-statistic has an F-distribution with I - 1 and I(M - 2) degrees of freedom if the null hypothesis H_0 is true. Thus, the test would be

$$\begin{cases} \text{Reject } H_0 \text{ if } F > F_{\alpha, I-1, I(M-2)} \\ \text{Do not reject } H_0 \text{ if } F \leq F_{\alpha, I-1, I(M-2)} \end{cases}$$

where $F_{\alpha, I-1, I(M-2)}$ denotes the upper α point of the F-distribution with I - 1 and I(M - 2) degrees of freedom.

Although it is important to examine each particular NRFB problem individually, it would be desirable to conduct, for the K NRFB problems, one overall test of the hypotheses:

$$H_0: \begin{cases} p_{11} = \cdots = p_{I1} \\ p_{12} = \cdots = p_{I2} \\ \vdots \\ p_{1K} = \cdots = p_{IK} \end{cases}$$

$$H_1: \quad H_0 \text{ is not true}$$

The use of external estimates $\hat{\sigma}^2_{ik}$ precludes a multivariate test because not enough information is available to permit calculation of the cross-product terms required in the matrix due to experimental error. [See

Winer (1971)]. A reasonable alternative is to analyze the data as arising from a repeated measures design. A diagrammatic view of the situation and the accompanying analysis of variance are given in Figure A1.

As can be seen from the figure, the usual repeated measures design does not permit a test of between-subject variation, which is usually not of interest. In the current situation, however, this variation is of interest and may be tested against the pooled error mean square

$$MSE = \frac{1}{IK} \sum_{i}\sum_{k} \hat{\sigma}^2_{ik} .$$

It should be noted that for a given subject, each $\hat{\sigma}^2_{ik}$ $(k = 1,\ldots,K)$ is based on the same M data points. Thus, MSE realistically has $I(M - 2)$ rather than $IK(M - 2)$ degrees of freedom.

From Figure A1 it can be seen that hypothesis H_0 (that the "correction" procedure is valid) is equivalent to $\sigma_S^2 = 0$ and $\sigma_R^2 = 0$. Thus, to test H_0, MSB and MSR must be treated against MSE. However, since it is desired to place H_0 in jeopardy, it is not enough just to set up a test with a given Type I error. A relatively small Type II error or large power (i.e., a good chance of detecting real differences) within reasonable limitations on number of subjects and problems is also required.

C. STATISTICAL POWER

If the experimental situation outlined in the previous sections is reviewed, it will be seen that departures from the null hypothesis should be reflected primarily in σ_R^2 (which measures the differential performance of subjects on problems), rather than in σ_S^2 (which measures between-subject

variability averaged over all problems). Thus, if the null hypothesis is false, it is reasonable to assume that σ_R^2 will be relatively large compared to σ_S^2. In view of this situation, the examination of power was focused on the test of $\sigma_R^2 = 0$.

It can be shown [Scheffe (1959), Winer (1971)] that the statistic $F = MSR/MSE$ is distributed as

$$[1 + (\sigma_R^2/\sigma^2)]F_{(I-1)(K-1),\ I(M-2)}$$

which reduces to $F_{(I-1)(K-1),\ I(M-2)}$ if H_0 is true. Thus, the power of this test depends on the size of σ_R^2 relative to σ^2, on the values of I (the number of subjects), K (the number of NRFB problems), and M (the number of RFB problems), and, of course, on the significance level α selected for the test. From a practical standpoint, it is desirable that the test be able to detect, with reasonable certainty, values of $\sigma_R^2 \geq \sigma^2$.

Figures A2 and A3 exhibit, for $\alpha = .05$ and $\alpha = .10$, power curves for different values of $\Theta = \sigma_R^2/\sigma^2$, I, K, and M. Based on these and similar curves, it was decided that values of I = 25, K = 8, and M = 10 provided minimum requirements.

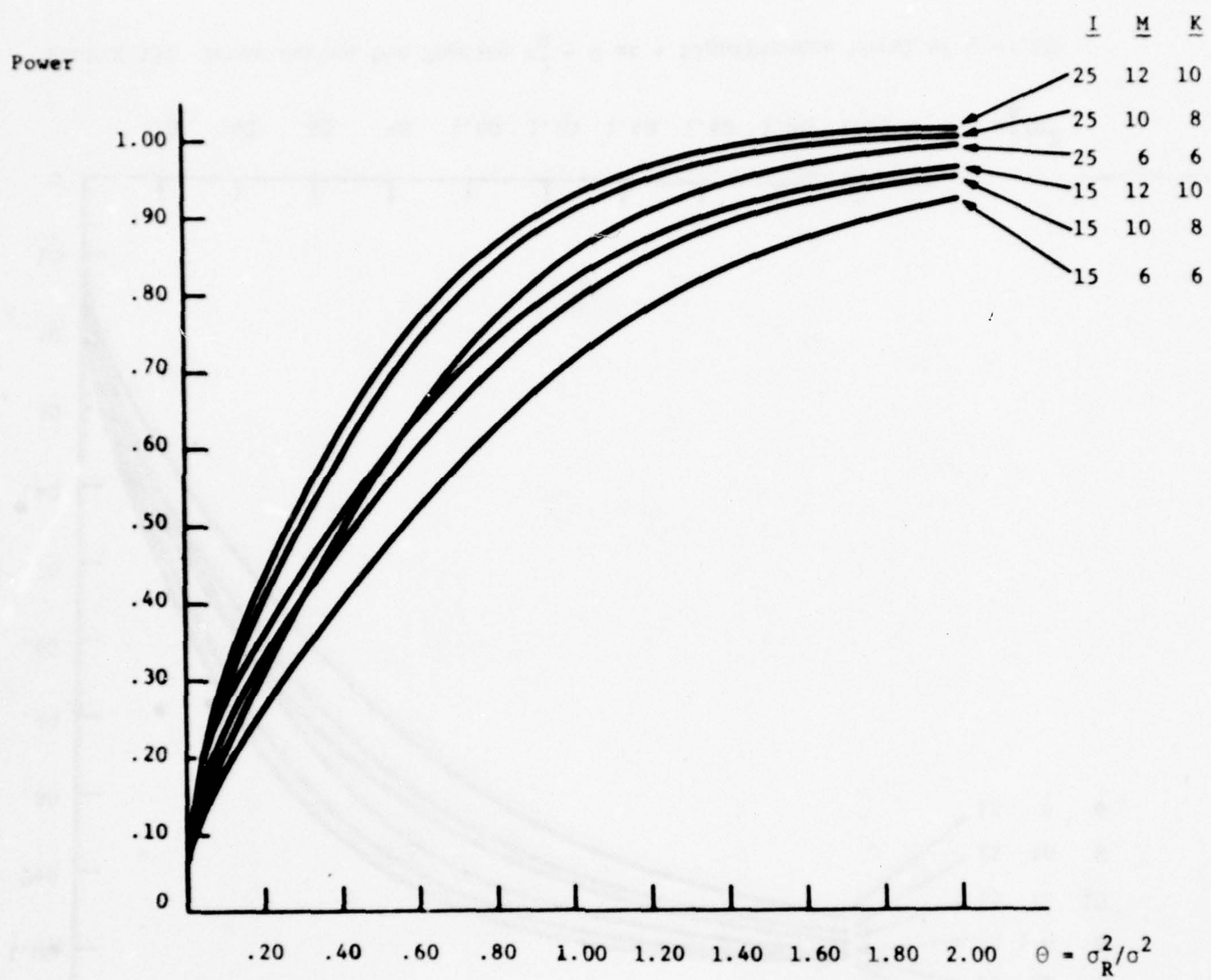


FIGURE A2: Power Curves for Testing $\sigma_R^2 = 0$ at a Significance Level of $\alpha = .05$

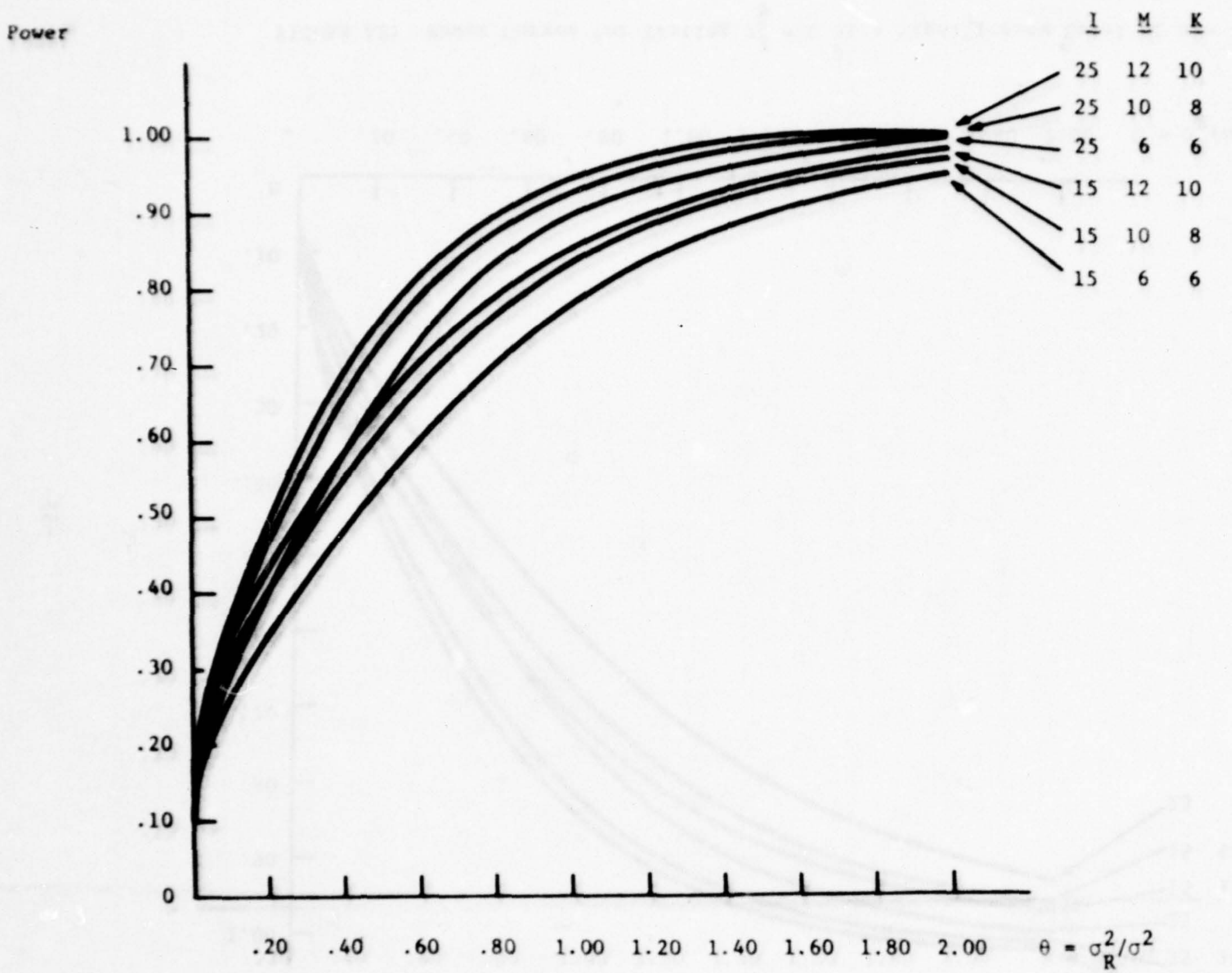


FIGURE A3: Power Curves for Testing $\sigma_R^2 = 0$ at a Significance Level of $\alpha = .10$

APPENDIX B: MATERIALS USED IN THE EXPERIMENT

The following pages comprise the set of materials used in the main experiment. These materials consist of an introductory section, examples, thirty-one (fourteen RFB, nine NRFB, eight almanac) problem situations, and a questionnaire.

DESMATICS, INC.

P. O. Box 618
State College, Pa. 16801
Phone: (814) 238-9621

DECISION-MAKING STUDY: INTRODUCTION

Purpose of Study

The purpose of this study is to provide information which can be used in development of improved decision-making methods. The study is being conducted by Desmatics, Inc. for the U.S. Army Research Institute for the Behavioral and Social Sciences under contract DAHC19-76-C-0045. The results will be published in a report at the end of the contract later this year. An abstract of the report will be published in a future issue of "Government Report Announcements and Index," a copy of which may be found in most large libraries.

Summary of Study

This study is concerned with probability, which is nothing more than the chance that something will or will not happen. There are many situations in which the term "probability" is used. For example, weather reports regularly contain statements such as "...the probability of rain is 40% tonight and 70% tomorrow..."

Suppose that a forecast of the probability of rain in the local vicinity is to be made for May 29, 1977. Assume that historical weather data indicates that for all the May 29 dates in previous years, there was rain on half of them and no rain on the other half. Based solely on this historical data, it would be reasonable to make an initial forecast for May 29, 1977, which

stated "....the probability of rain is 50%...."

However, as the May 29 date approaches, new information may change this forecast. For example, suppose that on May 28 there were rain clouds in the west, a falling barometer, and a change of wind direction so that prevailing winds were from the west. Most weather forecasters (and most other people) would feel that the initial forecast of 50% probability should be increased.

In most situations a probability cannot be precisely calculated--only estimated. Of course, probability estimates made by different people often vary. This study is aimed at determining the nature and extent of this variability in selected situations.

Today you will be presented with a number of problem situations. In each one you will be asked to estimate the probability (that is, the chance) of some specific thing happening. In making your estimates, try to have them reflect your judgment as accurately as possible, even if you find it difficult to decide on an estimate.

Some Guidelines for Making Estimates

When you are estimating probabilities, bear in mind that probability may range from 0% (for something which will <u>never</u> occur) to 100% (for something which is certain to occur). Also, it may be helpful to begin by mentally placing your estimate "in the right ball park." For example, in a certain situation you might feel that your estimate should be somewhere between 65% and 80%. Then "zero in" on your estimate by selecting a single value within this range.

Please try to make your final estimate so that you feel it is an accurate reflection of your judgment. For example, if you feel the probability is somewhere between 70% and 75%, do not estimate 70% or 75% solely because

they are nice round numbers. Instead ask yourself if perhaps a better estimate might be 71% or 72%. When you have made your decision, write down your estimate as a percentage between 0% and 100%.

Overview of Today's Session

During today's session you will be asked to estimate probabilities in a few simple situations which will be explained in a moment. At the end of the session, we would like you to answer a short questionnaire.

You will be given a set of materials which has an identification number on each page. This number will be used solely to correlate the various sections. To insure confidentiality, please DO NOT write your name on any page.

Your participation in this study will be limited to this one session. There will be no follow-ups of any kind.

DECISION-MAKING STUDY: EXAMPLES

General Description

The following pages contain three examples similar to those for which you will be asked to make probability estimates. In the first example you are asked to estimate a probability for a situation with which you have some general familiarity.

In the second and third examples, an initial set of conditions suggests that the required probability estimate is 50%. Then additional information is presented. Some of this information may tend to suggest an increase in the probability estimate, while some may suggest a decrease or no change at all. Your task is to evaluate all of the information and judge for yourself what the net effect would be.

Remember, you are to express your probability estimate as a percentage between 0% and 100%.

EXAMPLE A

There were over three million births in the United States in 1974.

Question: What would you estimate the probability to be that a baby born in the United States in 1974 was born east of the Mississippi River?

Answer: ______________________________ %

Note: This problem has to do with your estimate of whether in 1974 more babies were born east of the Mississippi than west of it. If you think so, your estimate should be greater than 50%. If you think there were less born east of the Mississippi your estimate should be less than 50%. Of course, if you believe an *equal number of babies were* born on each side of the Mississippi, your probability estimate should be 50%.

EXAMPLE B

At the teller's window of local banks there are various types of transactions including checking deposits, savings deposits and withdrawals, U.S. Savings Bonds purchases, traveller's check purchases, loan payments, check cashing, making change, payment of utility bills, and various combinations of these and other transactions. The most common type of transaction is cashing a check drawn on one's own account. In fact, local banks have found that 50% of all teller transactions involve cashing such a check.

Thus, it is reasonable to estimate that the probability is 50% that a *teller transaction* in a local bank will involve cashing a check on one's own account.

However such transactions are not equally distributed with respect to age and sex of the customer or day of the month. Some types of customers are more likely to cash their own checks than others, and particularly so at certain times and at certain offices.

Assume you are in line at a teller window in one of the bank offices within one block of campus. Suppose it is a Friday morning early in the month, when most checking accounts usually have a plus balance.

The man just ahead of you in line appears to be of college age. As far as you can tell, he is holding only one item, a check bearing the name of this bank. However, you cannot tell how it is made out or endorsed.

Question: What would you estimate the probability to be that the person ahead of you is cashing a check on his own account?

Summary

Initial probability estimate: 50%

Information:

(1) Bank near campus
(2) Friday morning early in month
(3) College age man
(4) Check on this bank

Answer: ______________________ %

Note: If you feel that the time, place and appearance of the person tends to suggest that it is a student cashing his own check, you should revise the probability estimate upward. On the other hand, if you feel the information tends to suggest it is not a student cashing his own check, you should revise the probability estimate downward. The degree to which you revise the estimate will be based on your own feelings about the information.

EXAMPLE C

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 75% blue chips and 25% red chips
Bag "R" contains 25% blue chips and 75% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "B" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is blue. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is blue. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is red.

Question: What would you estimate to be the probability that the selected bag is bag "B"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is blue
(2) Second chip is blue
(3) Third chip is red

Answer:_____________%

Note:
Since more blue chips were drawn from the bag, it is reasonable to estimate that the probability that bag "B" was chosen is greater than 50% because that bag contains more blue chips than bag "R" does.

Ident. No. ______

DECISION-MAKING STUDY: PROBLEM SITUATIONS

The following pages contain thirty-one problem situations which are similar to the three examples just presented. For each of these situations you will be asked to estimate a probability. These estimates will serve as data for analysis of the variability in probability estimates made by different people.

The first fourteen of the problem situations are of the "bag-and-chip" variety which you encountered in Example C. Although each bag-and-chip problem is similar in form, you must pay particular attention to (1) the color composition of chips in each of the two bags, (2) the results of the draws from the chosen bag, and (3) whether it is bag "B" or bag "R" that is referred to in the question.

Ident. No. ______

Problem Situation #1

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 66% blue chips and 34% red chips
Bag "R" contains 34% blue chips and 66% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "R" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is red.

Question: What would you estimate to be the probability that the selected bag is bag "R"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is red

Answer:____________%

Ident. No. ______

Problem Situation #2

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 75% blue chips and 25% red chips
Bag "R" contains 25% blue chips and 75% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "R" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is blue. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is blue. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is blue. The chip is returned to the bag, the contents are remixed and a fourth chip is drawn. This chip is red. The chip is returned to the bag, the contents are remixed and a fifth chip is drawn. This chip is blue.

Question: What would you estimate to be the probability that the selected bag is bag "R"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is blue
(2) Second chip is blue
(3) Third chip is blue
(4) Fourth chip is red
(5) Fifth chip is blue

Answer:____________%

Ident. No. ______

Problem Situation #3

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 70% blue chips and 30% red chips
Bag "R" contains 30% blue chips and 70% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "R" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is blue. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is blue. The chip is returned to the bag, the contents are remixed and a fourth chip is drawn. This chip is blue.

Question: *What would you estimate to be the probability that the selected bag is bag "R"?*

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is blue
(3) Third chip is blue
(4) Fourth chip is blue

Answer: ____________%

Ident. No. ______

Problem Situation #4

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 52% blue chips and 48% red chips
Bag "R" contains 48% blue chips and 52% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "R" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is blue. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is blue.

Question: What would you estimate to be the probability that the selected bag is bag "R"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is blue
(2) Second chip is blue

Answer:___________%

Ident. No. ______

Problem Situation #5

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 78% blue chips and 22% red chips
Bag "R" contains 22% blue chips and 78% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "B" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is blue. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is red. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is red.

Question: What would you estimate to be the probability that the selected bag is bag "B"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is blue
(2) Second chip is red
(3) Third chip is red

Answer: __________ %

Ident. No. ______

Problem Situation #6

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 70% blue chips and 30% red chips
Bag "R" contains 30% blue chips and 70% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "B" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is blue. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is blue. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is red. The chip is returned to the bag, the contents are remixed and a fourth chip is drawn. This chip is blue. The chip is returned to the bag, the contents are remixed and a fifth chip is drawn. This chip is blue.

Question: What would you estimate to be the probability that the selected bag is bag "B"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is blue
(2) Second chip is blue
(3) Third chip is red
(4) Fourth chip is blue
(5) Fifth chip is blue

Answer:__________%

Ident. No. ______

Problem Situation #7

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 67% blue chips and 33% red chips
Bag "R" contains 33% blue chips and 67% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "B" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is red. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is red.

Question: What would you estimate to be the probability that the selected bag is bag "B"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is red
(3) Third chip is red

Answer: ____________%

Ident. No. ______

Problem Situation #8

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 54% blue chips and 46% red chips
Bag "R" contains 46% blue chips and 54% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "R" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is red. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is blue.

Question: What would you estimate to be the probability that the selected bag is bag "R"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is red
(3) Third chip is blue

Answer:____________%

Ident. No. ______

Problem Situation #9

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 89% blue chips and 11% red chips
Bag "R" contains 11% blue chips and 89% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "B" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of tha selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is red.

Question: What would you estimate to be the probability that the selected bag is bag "B"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is red

Answer:____________%

Ident. No. ______

Problem Situation #10

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 85% blue chips and 15% red chips
Bag "R" contains 15% blue chips and 85% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "B" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is blue. The chip is returned to the bag, the contants are remixed and a third chip is drawn. This chip is red. The chip is returned to the bag, the contents are remixed and a fourth chip is drawn. This chip is blue. The chip is returned to the bag, the contents are remixed and a fifth chip is drawn. This chip is blue.

Question: What would you estimate to be the probability that the selected bag is bag "B"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is blue
(3) Third chip is red
(4) Fourth chip is blue
(5) Fifth chip is blue

Answer:____________%

Ident. No. ______

Problem Situation #11

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 90% blue chips and 10% red chips
Bag "R" contains 10% blue chips and 90% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "R" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is red. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is blue. The chip is returned to the bag, the contents are remixed and a fourth chip is drawn. This chip is red.

Question: What would you estimate to be the probability that the selected bag is bag "R"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is red
(3) Third chip is blue
(4) Fourth chip is red

Answer: __________%

Ident. No. ______

Problem Situation #12

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 59% blue chips and 41% red chips
Bag "R" contains 41% blue chips and 59% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "B" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is blue. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is red. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is red. The chip is returned to the bag, the contents are remixed and a fourth chip is drawn. This chip is red.

Question: What would you estimate to be the probability that the selected bag is bag "B"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is blue
(2) Second chip is red
(3) Third chip is red
(4) Fourth chip is red

Answer:____________%

Ident. No. ______

Problem Situation #13

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 65% blue chips and 35% red chips
Bag "R" contains 35% blue chips and 65% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "R" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is red. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is red.

Question: What would you estimate to be the probability that the selected bag is bag "R"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is red
(3) Third chip is red

Answer:____________%

Ident. No. _____

Problem Situation #14

Assume that there are two large identical bags, "B" and "R", each filled with a large number of poker chips. Although each bag contains the same number of chips, there is one difference:

Bag "B" contains 67% blue chips and 33% red chips
Bag "R" contains 33% blue chips and 67% red chips

Based on the toss of a coin, one of the bags is selected. The other bag is set aside. At this point, since each bag is equally likely to have been selected by the coin toss, it is reasonable to estimate that the probability is 50% that bag "R" was chosen.

Assume that a person who is securely blindfolded draws chips from the bag. Before a draw is made, the contents of the selected bag are thoroughly mixed. Then one chip is drawn. The chip is red. The chip is returned to the selected bag, the contents are again mixed and a second chip is drawn. This second chip is red. The chip is returned to the bag, the contents are remixed and a third chip is drawn. This chip is blue.

Question: What would you estimate to be the probability that the selected bag is bag "R"?

Summary

Initial probability estimate: 50%

Information:

(1) First chip is red
(2) Second chip is red
(3) Third chip is blue

Answer: ____________%

Ident. No. ______

Problem Situation #15

Assume that GM sells the same number of 1978 trucks that are sold by non-GM manufacturers. Thus, if in a conversation with a stranger, you found out that he had purchased a new 1978 truck, it would be reasonable to estimate that the probability is 50% that his truck was a GM.

Suppose that, at the beginning of the next model year, GM sustains a massive strike for three months during which no 1979 trucks are produced, while all non-GM truck manufacturers avoid strikes. Further suppose that, midway through the model year, a Ralph Nader-type group issues a report that 1979 GM trucks have serious safety defects. In addition, assume that although 1979 GM trucks increased in price by an average of $190 over the previous year's price, non-GM trucks increased in price by an average of $345.

Sometime in 1980 you meet another stranger whom you find has purchased a new 1979 truck.

Question: What would you estimate to be the probability that it was a GM truck?

Summary

Initial probability estimate: 50%

Information:

(1) GM strike
(2) GM safety questioned
(3) Smaller GM price increase

Answer: ______________________________%

Ident. No. ______

Problem Situation #16

You are considering taking a particular elective course next term. When you first asked about it a few weeks ago, the head of the department offering the course told you that it might or might not be given, and that there was a 50% probability that it would be given next term.

A few weeks later you pass the department head in the hall. He tells you that the instructor who taught the course last year will be on leave next term. You also see a notice on the bulletin board requesting persons interested in this particular course to sign up. The notice is about a week old and there are only three signatures on the notice. In addition, you read in the paper that because of budget problems, some elective courses may not be offered next term, but postponed until later.

Question: In light of these factors, what would you estimate to be the probability that the elective course you are interested in will be given next term?

Summary

Initial probability estimate: 50%

Information:

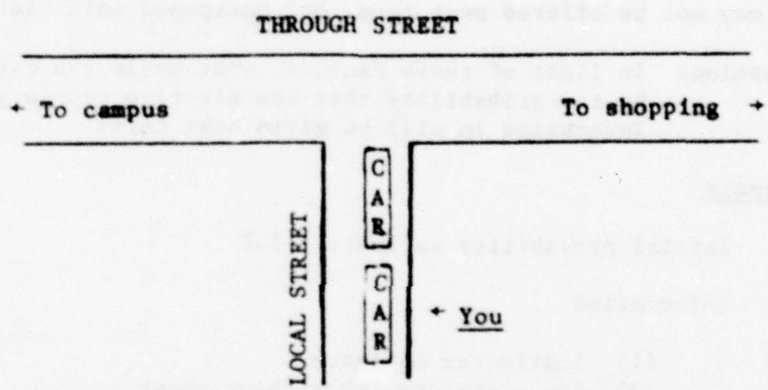
(1) Instructor on leave
(2) Few signatures on sign-up sheet
(3) Postponement policy announced

Answer: ______________________%

Ident. No. ______

Problem Situation #17

There is a street corner in town where a "local" street intersects a "through" street in a "T". The local street does not continue at this point, so only turns onto the through street are possible for cars arriving at this intersection from the local street. Left turns lead to the university campus, while right turns lead to shopping centers. During an average weekday, as many cars arriving at this intersection from the local street turn left as turn right. Thus, it is reasonable to estimate that the probability is 50% that a car arriving at this intersection will turn left.



Suppose you approach that intersection from the local street at 10:45 a.m. on a weekday when university classes are in session. You note that the car ahead of you has a faculty/staff parking sticker on the rear bumper. A woman driver is the sole occupant.

Question: What would you estimate to be the probability that this particular car will turn left?

Summary

Initial probability estimate: 50%

Information:

(1) 10:45 a.m.
(2) university classes in session
(3) faculty/staff sticker
(4) woman driver

Answer: ______________________________%

Ident. No. ______

Problem Situation #18

Although Springfield is the capital of Illinois, the population of that state is concentrated around Chicago, which is located approximately 200 miles northeast of Springfield. In fact, the population of the Chicago metropolitan area is almost exactly half the total population of the state of Illinois. Thus, it is reasonable to estimate that the probability is 50% that a person from Illinois is from the Chicago metropolitan area.

Suppose that on a nonstop plane from Washington, D. C. to Chicago you meet a stranger who says he lives in Illinois. He mentions that he works for the State of Illinois. He also mentions that he is a sports fan and attends many of the Chicago baseball games. When you arrive in Chicago and retrieve your carry-on luggage you note that his bag has several Springfield, Ill., stickers. He is also carrying parcels that are labelled "Illinois Department of Agriculture."

Question: What would you estimate to be the probability that this stranger resides in an area of Illinois other than the Chicago metropolitan area?

Summary

Initial probability estimate: 50%

Information:

(1) Works for State of Illinois
(2) Attends Chicago sports events
(3) Bag has Springfield stickers
(4) Parcels say "Illinois Department of Agriculture"

Answer: ______________________________%

Ident. No. ______

Problem Situation #19

In New York harbor, most of the passenger facilities are in Manhattan. Brooklyn, on the other hand, has the majority of the large cranes which facilitate unloading heavy deck cargo. Normally about half the ships which enter New York harbor from the Atlantic dock in Manhattan. Most of the remainder dock in New Jersey or in Brooklyn. Thus it is reasonable to estimate that the probability is 50% that a ship entering New York harbor from the Atlantic will dock in Manhattan.

Suppose you have read in the paper that ships docking in New Jersey are currently experiencing extended delays in unloading due to a workers' slowdown. Today, as you cross the bridge at the New York harbor entrance, you see a large cargo ship entering. This ship appears to be loaded with heavy deck cargo.

Question: What would you estimate to be the probability that this ship will dock in Manhattan?

Summary

Initial probability estimate: 50%

Information:

(1) New Jersey slowdown
(2) Cargo ship
(3) Heavy deck cargo

Answer: ______________________________%

Ident. No. ______

Problem Situation #20

On an average day the local A&P supermarkets sell the same amount of Pepsi and Coke---some in cans, some in throwaways, and some in returnable bottles. Thus, *it is reasonable to estimate* that the probability is 50% that a local A&P customer who buys one of the two leading colas will buy Pepsi.

This week the local A&P is having a sale in which 8-packs of Pepsi returnable 16-ounce bottles are on sale for 50¢ less per carton than Coke. Furthermore, the store has a special display of Pepsi products near the entrance. However, a nearby supermarket competitor is featuring a 50¢ per carton reduction on 8-packs of Coke, announced in a large window sign and prominently featured in newspaper ads.

Question: In light of this information, what would you estimate to be the probability that a local A&P shopper who buys one of the leading colas this week buys Pepsi instead of Coke?

Summary

Initial probability estimate: 50%

Information:

(1) A&P Pepsi sale
(2) A&P Pepsi display
(3) Competitor Coke sale

Answer: ______________________%

Ident. No. ______

Problem Situation #21

There were almost 80 million votes cast in the U.S. presidential election in November 1976. Jimmy Carter received approximately 40 million votes. Thus, it is reasonable to estimate that the probability is 50% that a person who voted in the 1976 presidential election voted for Carter. However, Carter received 56 more electoral votes than Ford and was elected. Carter ran stronger in some regions than others and was more popular with some elements of the electorate than others.

Suppose that on a trip to Georgia you meet a black woman who tells you she is 53 years old and has lived in Georgia all her life. You also learn from her that she is married and works in a textile mill. She mentions that she voted in the 1976 presidential election.

Question: What would you estimate to be the probability that this woman voted for Jimmy Carter?

Summary

Initial probability estimate: 50%

Information:

(1) Black woman
(2) 53 years old
(3) Lifetime Georgia resident
(4) Married
(6) Works in textile mill

Answer: ______________________________%

Ident. No. ______

Problem Situation #22

Roughly half the babies born in this country are girls, so lacking other information it is reasonable to estimate that the probability is 50% that a particular birth in the local hospital will be a girl.

Mrs. Jones is nearing the time for her baby to be born. Her doctor thinks it may be a girl because its heartbeat is somewhat faster than the average for boys. The nurse, noting that the Jones have three boys, thinks it will be a boy because she feels that boys run in their family. Mr. Jones, a casual dice player, feels it will be a girl because "it is time for the run to change," thus helping to even the trend.

Question: What would you estimate to be the probability that Mrs. Jones' baby will be a girl?

<u>Summary</u>

Initial probability estimate: 50%

Information:

(1) Faster heartbeat
(2) Nurse's opinion that boys run in family
(3) Father's opinion that a girl would help even trend

Answer: ______________________________%

Ident. No. ______

Problem Situation #23

Assume you are in class on the fourth floor of a seven story classroom building, so there are three floors above this floor and three floors below. You leave your class a half hour early because of a dental appointment and go to the only public elevator. The elevator is not there, so you push the button. On the average, in this type of situation, the elevator is as likely to have been left by the previous user on a floor above as on a floor below. Thus, it is reasonable to estimate that the probability is 50% that the elevator will come up from below in response to your call.

However, in this case you realize that you are not taking into account all of the circumstances. In this building the seventh floor is primarily instructors' offices, the sixth floor is mostly research laboratories, and the remaining floors are all classrooms. Furthermore, half of the third floor classrooms are being painted. In addition, it is early in the morning since it was your first period class that you left.

Question: Under these circumstances what would you estimate to be the probability that the elevator will come up from below in response to your call?

Summary

Initial probability estimate: 50%

Information:

(1) Seventh floor offices
(2) Sixth floor labs
(3) Other floors are classrooms
(4) Half of third floor rooms being painted
(5) Morning, middle of first period

Answer: ____________________________%

Ident. No. ______

Problem Situation #24

New York state is the largest patron of the arts, leading all other states in 1975 in terms of state appropriations for the arts.

Question: What would you estimate to be the probability that a dollar appropriated for the arts by one of the states in 1975 came from a New York state appropriation?

Answer: ________________________________%

Ident. No. ______

Problem Situation #25

Philadelphia is the largest city in Pennsylvania in population, according to the 1970 census.

Question: What would you estimate to be the probability that a Pennsylvania resident in 1970 was from Philadelphia?

Answer: ______________________________%

AD-A077 994 DESMATICS INC STATE COLLEGE PA F/G 12/1
AN APPROACH FOR PROVIDING MORE ACCURATE PROBABILITY ASSESSMENTS--ETC(U)
MAY 78 D E SMITH , R L GARDNER DAHC19-76-C-0045
UNCLASSIFIED ARI-TR-78-B7 NL
2 OF 2
AD-
A077994

END
DATE
FILMED
1-80
DDC

Ident. No. ______

Problem Situation #26

There were over 16,000 deaths in the U. S. in 1974 due to falls.

Question: What would you estimate to be the probability that a person who died in a fall in the U. S. in 1974 was over 54 years old?

Answer: ____________________________%

Ident. No. ______

Problem Situation #27

According to Federal figures there were almost 133 million motor vehicles registered in the U. S. in 1975.

Question: What would you estimate to be the probability that a motor vehicle in the U. S. was registered in Pennsylvania in 1975?

Answer: ______________________________%

Ident. No. ______

Problem Situation #28

FBI figures for 1975 report that there were 116,000 arrests in the U.S. that year for robbery.

Question: What would you estimate to be the probability that a 1975 U. S. arrest for robbery was a man rather than a woman?

Answer: ______________________%

Ident. No. _____

Problem Situation #29

Automobiles are the single most common cause of accidental death in the U. S.

Question: What would you estimate to be the probability that a person who died in an accident in the U.S. in 1974 was killed in an auto accident?

Answer: ____________________________%

Ident. No. ______

Problem Situation #30

There were over 6.7 million new passenger cars produced in the U. S. in 1975.

Question: What would you estimate to be the probability that a 1975 U. S.-manufactured passenger car was a product of General Motors?

Answer: ______________________________%

Ident. No. ______

Problem Situation #31

There are a number of departments in the U. S. federal government.

Question: What would you estimate to be the probability that a dollar of Federal tax money budgeted for expenditure in fiscal year 1976 was allocated for expenditure by the Department of Health Education and Welfare rather than some other department of the federal government?

Answer: ____________________________%

****

THIS IS THE LAST PROBLEM SITUATION. AT THIS POINT, PLEASE CHECK BACK TO SEE THAT YOU HAVE PROVIDED ANSWERS FOR ALL 31 PROBLEM SITUATIONS.

AFTER YOU HAVE DONE THIS, PLEASE COMPLETE THE SHORT QUESTIONNAIRE ON THE FOLLOWING TWO PAGES.

Ident. No. _____

SUPPLEMENTARY INFORMATION

1. Some of the problems you have just completed established an initial 50% probability estimate which you were asked to revise on the basis of additional information. Please indicate below how valid you think the initial 50% probability estimate was in each case. Then indicate how sure you feel about the probability estimates you made. Finally, indicate how realistic you feel each situation was. Place check marks to indicate your evaluation.

PROBLEM SITUATION	How Was Initial 50% Probability Estimate?			How Sure Were You of Your Estimates?			How Realistic Were Situations?		
	Seemed Okay	Off Slightly	Seriously Off	Quite Accurate	Only Fairly Accurate	Sheer Guess	Fairly Accurate	A Bit Contrived	Completely Hokey
Bag and Chip Problems									
GM Trucks									
Elective Course									
Street Intersection									
Illinois Resident									
New York Harbor									
Pepsi vs. Coke									
Presidential Election									
Mrs. Jones' Baby									
Elevator									
Problems with No Initial Probability Estimate									

(Reduced from full-page size.)

Ident. No. ______

2. Do you think there might be a way to <u>calculate</u> the bag and chip probabilities? ______________

   If so, how?

3. Have you ever, in school or elsewhere, encountered situations requiring you to:

   a. Make probability estimates? ______________

   b. Revise probability estimates on basis of subsequent information? ______________

4. How many years of education have you completed beyond high school (including the current year)? ______________

5. Major field ______________

6 Minor field ______________

7. Approximate overall grade point average ______________

8. Indicate the number of credits (if any) completed or currently scheduled in:

   Probability __________ Statistics __________ Decision Theory __________

8. Age __________

9. Sex __________

10. Comments: